The British Journal for the Philosophy of Science

VOLUME I NOVEMBER, 1950 No. 3

INDETERMINISM IN QUANTUM PHYSICS AND IN CLASSICAL PHYSICS

PART II*

8

We now proceed to the discussion of the theorem which asserts that the computing part of a predictor must be incapable of producing an accurate answer to every question concerning its own future physical states, whatever the amount of information it has received about its own past or present states. (As a kind of explanation of this fact, it will be shown that however full this information may be it will always be incomplete in an important way.) Throughout this section, we shall confine our investigation to the calculator part of the predictor; and it does not signify whether we assume that the calculator part receives its information from its own measuring parts or from some other (outside) source.

In view of the rôle played by this theorem in our whole argumentation, I shall discuss it at some length. My discussion will consider three arguments, in some respects similar and not unconnected, but in other respects independent. The three arguments may be briefly described, for easier reference, (I) as a variant of the so-called '*paradox of Tristram Shandy*'; (II) as a physical application of so-called '*Goedelian sentences*'; and (III) as a utilisation of what I usually call the '*Oedipus effect*'.[1]

---

* Part I appeared in volume 1, no. 2 of this *Journal*.

[1] Only the last of these three was discussed in my original paper. See note 1 on p. 117, above (volume 1, no. 2 of this *Journal*).

(I) The '*Tristram Shandy paradox*' is a name usually given to the following difficulty. Tristram Shandy tries to write a very full story of his own life, spending more time on the description of the details of every event than the time it took him to live through it. Thus his autobiography, instead of approaching a state when it may be called reasonably up to date, must become more and more hopelessly out of date the longer he can work on it, i.e. the longer he lives. The so-called paradox now arises as follows. In spite of the fact that the gap, existing at any moment between the last event described in his story and the period of writing, increases (towards infinity) as his span of life is prolonged, the gap would, nevertheless, disappear, if only he could go on living for ever ; for then he could ' complete ' his task in the sense that every event of his life, however late, would be covered by some part of his autobiography.

We shall be concerned here with a somewhat different aspect of this story, viz, with the following. Even if Tristram Shandy (or a calculator *C*) is as quick as we choose in the full description of his history, he must be incapable, nevertheless, of bringing it completely up to date. For assume that at the time t_2, he has completed his story up to the moment t_1, and now tries to describe the last period of his life, i.e. the period between t_1 and t_2. The period of time it takes him to do this will be the period between t_2 and t_3 ; that is to say, when he has completed his description up to t_2, he has still the period between t_2 and t_3 to describe. A full description of this period involves the full description of what he wrote between t_2 and t_3, in other words, even if he neglects all other events of his life during that period, he must at least describe every word he wrote between t_2 and t_3. When this is completed, the time will be t_4, and he has still before him the task of describing his doings between t_3 and t_4 ; and so on.

There is, of course, no paradox here. One may perhaps even say that, if Tristram Shandy continually improves his methods of description or the speed of his writing, he may bring his autobiography more and more nearly up to date, and that there is no limit to this procedure. Thus, in so far as the *description of the past* is concerned, there is no serious difficulty here. But our consideration leads to another which shows that very serious difficulties arise in connection with a *prediction of the future*.

To see this, we need not consider a complete predictor, but merely its calculator part, the calculator *C*. We assume that some standardised information may be fed into *C*—say, in form of a tape with punched holes—and that C has been set so as to produce a prediction for the instant t_i of the state of any physical system about whose state at t_0 we inform C. The prediction, we may assume, also takes the shape of a tape with holes ; so that C may be described as a punching mechanism which, if stimulated by a certain 'setting for t_i', followed [1] by the insertion of a punched tape, punches holes into another tape.

We further assume that C is endowed with a 'memory', i.e. a chamber in which results of its calculations—the final predictions produced—and the information received are stored ; the state of this memory at any time will be an important part of the state of the machine at that time. Another assumption is that the punching mechanism of C, once it embarks on a predicting task, runs on until it has completed its task ; then it stops.

Now we assume that at some instant of time t_1, C has received accurate and complete information about its state (an assumption whose possibility will be challenged and investigated more fully in our third argument—the one connected with the *Oedipus effect*) and that C has previously been set to predict its own state at a certain later instant, say, at t_5. Assume that, at t_2, C has succeeded in predicting the t_5 state of the particles making up its wheels, its punches, etc., and also that of the particles making up part of its 'memory', viz, that part which constitutes its memory of the information which it has received ; and that it now—at t_2—proceeds to describe the state of that part of its memory in which its predictions are stored. Assume further, that it proceeds at t_2 to describe that part of the prediction tape which it produced between t_1 and t_2. When this task is completed, the time will be t_3. At t_3, therefore, it may begin to describe that part of the tape which it produced between t_2 and t_3 ; when this is completed (and C is still faced with the task of describing the state of the tape produced between t_3 and t_4) the time will be t_4 ; etc.

It is perfectly conceivable that these various time intervals may be made shorter and shorter.[2] But since each of them must contain *some* description, to be described in the next interval, and since the assumption seems unavoidable that, for every actual physical machine, there is a maximum running speed, and consequently, a minimum length of time necessary for completing even the shortest description of which the machine is capable, we cannot assume that the series of time intervals which we have described can be actually converging. In other words, the machine will not complete its task, and will continue to punch holes, at any rate until t_5 has passed. Since its task will be completed when it has described that piece of tape which was punched just before, or at, t_5, it *may* stop at some instant shortly after t_5, but not before. But this means that *C*'s prediction of *C*'s own state at t_5 will be completed too late to be a real prediction. Or more generally, although *C* may describe its own t_i state in detail, it can never complete this description before t_i has passed ; it therefore can describe completely only its own past, but not its own future.

---

[1] By assuming that the time t_i for which C is to make its prediction was set prior to inserting the last information—that C is switched on, as it were, and perhaps already computing, but at any rate capable of taking account of the latest information we insert into it—we avoid certain unessential complications ; for this makes it in principle possible that our information is as highly up to date as it can be, since we do not need to interfere with C after supplying it with information. And this information *may* tell C that and when it has been set for t_i.

[2] This may be achieved, for example, by letting C run with constant acceleration (however small), if, for example, the length of tape l_{n+1} necessary for describing

I think that in this form,[1] the argument is not yet quite conclusive. For it is conceivable that a machine may be constructed which can formulate statements capable of describing themselves and *at the same time* other statements (for example, the one which immediately precedes it on the tape). Such a machine may perhaps use numbers as names of expressions of a certain shape ;[2] and it may then conclude its self-prediction by the following statement (occupying, say, the last 4 yards of the tape) : ' The last fifteen yards of the tape are taken up in this way : the first twelve yards by two equally long series of holes of shapes No. 17,623,412 and 34,216,311, and the remainder by a series of holes of shape No. 612,522,947.' And it may turn out that the last mentioned number is indeed the number of the shape of the statement just quoted, perhaps even in conjunction with other statements.

In spite of this objection, our argument may be made quite conclusive by introducing the postulate (which, however, seriously

---

the preceding piece of tape of the length l_n equals the latter plus a constant ; this constant can, we may imagine, be the space needed for the prefix ' The next piece of tape consists of holes representing the sentence quote ' . . . and the suffix ' unquote '. However, this method of describing the tape is not one of a standardised physical description of its state ; see below.

[1] In the form given, the argument is somewhat similar to one given by Prof G. Ryle in *The Concept of Mind* for a different purpose—the clarification of the elusiveness of the *Ego*.

[2] This is a method introduced bv Goedel ; cp. note 1 on p. 180 and note 2 on p. 183, below.

limits the generality of our consideration[1]) that the machine must retain, in its memory, traces not only of the final formulations of its predictions but also the more immediate partial results of its deductions. For we must remember that our machine is a calculator or deduction machine, and that, as such, it must carry out its considerations, at least at first, in some sort of standardised language or standardised symbolism (for example, consisting of sequences of figures) ; and we may perhaps assume that, when it tries to predict, from physical premises, the physical state of its memory, it will have to do so, in the first instance at least, by way of deducing a detailed physical description which must contain, at least,[2] descriptions of the physical shapes of every elementary sign contained in its memory (for example, of every digit constituting the numbers).

But for this kind of secondary description—i.e. for one which is formulated in the same language and which utilises the same kinds of physical symbols as it describes—the lemma holds that it must always be at least as long as the primary description which it describes.

This lemma is obvious ; for every elementary sign of the original description must be one of the physical entities to be described, and none of these descriptions can consist of less than one elementary sign (in fact, it must consist of more). But if this lemma is true,[3] then our argument will hold good for the more immediate partial results of the deductions computed by *C*. If these are to be retained by its memory (as we postulated), then we may confine our attention to them. And if *C* contains in addition special methods permitting it to transcribe such results into radically abbreviated (and possibly self-describing) statements, then we need not consider these at all.

From the lemma mentioned it also follows that our argument is valid independently of the order in which *C* describes the contents of its memory. More particularly, C does not need to describe these contents in historical order. Any description whatsoever will take up as much, or more, tape than the tape to be described ; and since the amount of tape which can be used before t_5 must be finite, the description of the t_5 state of the memory of C cannot be completed before t_5, whether the machine is fast or slow.

---

[1] Although at some instant of the machine's running, these partial results must be in untranscribed form (see below), and this instant may be made to coincide with t_5.

[2] *C* may ' know ' that in dealing with the effect of information on a computor, only the signs count and not the accidental variations in their shape ; and for this reason, it may give a somewhat summary description of them.

[3] Another way of proving this would be to consider the ' *amount of information contained in a message* ' (a cybernetic term). The second message which describes the first must contain at least the same amount of information as the first. It may be remarked that our lemma may be expressed in more general terms, referring to every fixed method of recording messages, and to the detailed physical description (recorded by the same method) of so recorded messages.

Taking this to be established, we find that it is impossible for C completely to predict the state of its tape for any future instant of time t_i. Now it might be argued further—but I am not sure whether validly—that the state of the tape at later instants will strongly depend on that at earlier instants ; and that since a complete prediction is impossible even for the nearest future, the amount of foreknowledge available to the machine of the state of its tape must be less extensive than we thought when we made the assumption that, at t_3, the machine had succeeded in describing part of the t_5 state of its tape.

There are further such assumptions which in the light of this consideration now appear highly questionable, especially the one that the machine could complete (at t_2) the prediction of its state, viz, the state of its wheels, etc., except the state of its memory. For there will be a one-one correlation between the various states of its wheels and those of the tape which is produced by these wheels ; and although the state of the tape may be looked upon as ' caused ' by that of the wheels, the dependence will hold in both directions. We can, therefore, consider the state of the wheels as depending on that of the tape ; we can, as it were, identify for the moment the machine C with its tape, so that the rest of the machine becomes the closest environment of C, i.e. that part of the environment which is most strongly dependent upon C. Now if it is impossible for C completely to predict its state (i.e. that of its tape) for *any* moment of its future, we may have to conclude that it will also be impossible for it to predict completely the state of its closest environment (i.e. the state of its other parts).

But this might seem, at first sight, to lead to a serious difficulty. For why should a machine C be incapable of predicting the future state of C if a very similar machine C^+ does not encounter any obstacles if faced with the same task—that of predicting C ? We assumed that any amount of information was supplied to C. Thus, it appears, C^+ cannot have more information than C. Furthermore, C need not ' know ' that the calculations which it performs refer to itself. It is simply given a set of initial conditions, and the task to deduce predictions. Why should C fail where C^+ succeeds ?

The answer to this question is simple but important. If we furnish C and C^+ with the same information (e.g. about C) and set them the same task, they will of course produce precisely the same predictions. Now we assume we have furnished only C^+ with this accurate and full information about C, and that C^+ alone produced predictions about C's t_5 state. No doubt these predictions will be true, and C will be at t_5 in the state predicted. But if C also is given that information, this will mean that we have interfered with C, and that its t_5 will be different from the previous case. Accordingly, the t_5 prediction (which both C^+ and C produce in conformity) will no longer be true. In other words, by informing C^+ alone, we get a true prediction, while by informing both, we obtain from them identical but false predictions. And this explains why C^+ can predict C, while C cannot predict C. The reason is that C^+ can be given information about C which is up to date, while if the same information is given to C, its up-to-dateness is destroyed.

This shows that the apparent difficulty in our view dissolves at closer inspection. At the same time, a new and perhaps more promising line of argument is opened—that of examining the effect of the information received by C on its capability of self-prediction. This argument will be taken up more fully below when trying to utilise the ' *Oedipus effect* '.

Against our whole Tristram Shandy argument, two serious objections may be raised. The first has been indicated ; it is the objection that we must introduce a special postulate, amounting in its effect to the exclusion of the possibility (or to an excuse for neglecting the possibility) that the machine produces statements referring to themselves. It seems likely (especially in view of certain results of Goedel) that such a postulate amounts to a serious limitation of the mathematical powers of the machine. The other and even more vital objection to our argument is this. Even if it is valid—does it establish what we wish to establish ? Is it not still conceivable that a machine, although incapable of giving a complete description of its future physical states, may answer every sufficiently definite question about its physical state at a future time t_i with ' yes ' or ' no ', before the instant t_i has passed ?

(II) The consideration of a physical application of *Goedelian sentences* will help us to answer this last question in the negative.

Goedel's famous and much discussed theorem (published in

1931)[1] established, among more general results, the fact that for every sufficiently formalised and sufficiently powerful[2] deductive system S_1 of true statements, there exist true mathematical formulae which can be completely expressed in the terms which are used by S_1, but which are neither demonstrable nor refutable in S_1. (If a demonstration can be found in S_1, then S_1 can be shown to be inconsistent.) The truth or falsity of any one of these ' Goedelian sentences ' can be demonstrated in a strengthened system S_2 (which may be constructed, for example, by adding the Goedelian sentence to the axioms of S_1). But S_2, again, can be shown to be similarly deficient, since new Goedelian sentences can be constructed which are undecidable in S_2 ; and so on.

Goedel further showed—and this is important for our purposes—that if g is a Goedelian sentence with respect to S_1 (that is to say, g is undecidable in S_1), then the true statement h which asserts that g is undecidable in S_1 is, in its turn,[3] also undecidable in S_1.

Now a formalised deductive system S_1 may be said to correspond to a calculating machine C ; and this fact raises hopes for utilising Goedel's results for our purposes. Nevertheless, there are difficulties. What we wish to construct is a physical prediction statement—one about a physical state of C—which cannot be decided by C. But the Goedelian statements are all statements of pure mathematics, and as such different from statements about concrete physical states of concrete physical machines. This is the reason why I was unable, for some time, to utilise the Goedelian method for my purposes.[4]

Ultimately, I found that the difficulty could be solved in a comparatively simple way. It is easy to construct a sentence which corresponds to what I have just described as ' the true statement *h* ' and which, at the same time, is formulated as an assertion about *C*'s future physical powers of reacting to a certain physical stimulus. If such a sentence is put as a question to *C*—that is to say, in such a way as to stimulate *C* to react, if it can, by signalling ' yes ' or ' no ' —than *C* will be unable to decide it before the decision has lost its predictive character. Since this fact follows nearly immediately from Goedel's results, I shall call this kind of undecidable question a ' near-Goedelian ' question.

---

[1] K. Goedel, ' Ueber formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I ', *Monatsh. Maths. Phys.*, 1931, **38**. See also J. B. Rosser, ' Extension of some Theorems of Goedel and Church ', *Journal of Symbolic Logic*, 1936, **1** ; A. Tarski, ' On Undecidable Statements in Enlarged Systems of Logic and the Concept of Truth ', *ibid.*, 1939, **4** (see esp. theorem 6.2) ; and A. M. Turing, ' On Computable Numbers, with an Application to the Entscheidungsproblem ', *Proc. of the London Mathem. Soc.*, ser. 2, 1937, **43**.

[2] The essential condition is that the system must contain arithmetic and such elementary concepts of the theory of numbers as ' Prime Number '.

[3] It is even a consequence of Goedel's (second) theorem that—provided no false statement is derivable in S_1—every true statement u which asserts the non-derivability in S_1 of some statement or other (e.g. of the statement ' $0 = 1$ ') must, in its turn, be undecidable in S_1.

[4] I dwelt at length, however, upon the close analogy between my problem and Goedel's in my original paper. Cp. note 1 on p. 117, above (volume **1**, no. 2 of this *Journal*).

But once we have these near-Goedelian questions, we may even construct a more concrete and specific physical question—a question whether *C* will react *within a given finite interval of time* to a certain physical stimulus. This second kind of question can also be shown to be answerable by *C* only after the event—when it has lost its predictive character. A question of this second and more specific kind is what I call a ' semi-Goedelian ' question. I shall discuss the two kinds of questions in turn ; and in order to show how they are formulated, I shall first describe the reactions of which *C* is capable.

We consider a calculator *C* capable of responding with ' yes ' whenever it finds that a statement q_i—the question which is put to *C* at the instant of time t_i—is derivable from the information J_i with which *C* is supplied at the instant t_i. When *C* finds out that the negation[1] of q_i is derivable from J_i, then it responds by signalling ' no ' ; and until it has decided q_i, in one way or another, it signals, on being asked q_i, ' I do not know yet.' (More detailed assumptions about the method of the working of *C* will be mentioned later.)

Now we can ask *C* at the time t_1 a question q_1 about its own future physical t_2 state. For example, q_1 may be the question ;[2] ' If you are asked at t_2 the question q_2 (and allowed any amount of time for an answer) will or will you not ultimately react by signalling "yes" or "no" to q_2?' It can be shown that, if q_2 happens to be a Goedelian sentence (with respect to J_2), then C cannot make any other reply to q_1 than 'I do not know yet'—at any rate, *not before the instant* t_2 has passed; and this will be the case however full the information may have been of which C was possessed at t_1—provided always this information J_1 was true (and therefore consistent), and that it formed part of J_2. (Whether or not C will in fact answer the question q_1 when it is too late for the answer to be acceptable as a real prediction will depend on C's construction; see below.) The fact that C cannot answer the 'near Goedelian' predictive question q_1 shows that C is unable to predict at least some of its own future reactions—some aspect of its own future physical t_2 state.

---

[1] Or, in case the question q_i is a conditional one, the conditional question whose consequence is negated. We may assume that *C* can be asked both questions at a time, as our formulation—' will you or will you not '—indicates.

[2] The question as formulated here may not be in the standardised form which is ' understandable ' to *C* ; the question q_1 may, for example, be inserted into *C* in form of a punched tape whose punches represent a number (see below) and it may be the number of the statement : ' If into the opening No. 2 of the physical system *C* (about which information has been previously inserted), a tape is inserted at t_2 with such and such holes, will certain wheels *W* of *C* be agitated ? ' Of course, the holes described would be those which represent the number of q_2, i.e. of a Goedelian sentence relative to J_2.

Note that the Goedelian question q_2 need not really be asked at t_2; all that happens is that C is asked the question q_1 which is a question about C's future physical reaction to a definite physical stimulus. Thus the near-Goedelian question q_1 is a physical question—unlike q_2 which is a Goedelian question of pure mathematics. But the fact that C cannot answer a near-Goedelian question such as q_1 is, in its turn, capable of purely mathematical demonstration, and thus a mathematical truism about all computers satisfying certain general conditions.

I now proceed to a fuller analysis of the situation.

As one of the simpler examples of a sentence which is Goedelian realtive to S_1 (i.e. undecidable within a sufficiently rich system S_1) we may choose [1] a statement g asserting that the sentence '$0 = 1$' is not derivable from the premises of S_1. (In other words, g asserts that S_1 is consistent.) As Goedel has shown, it is always possible to formulate in S_1 a purely arithmetical sentence g' which corresponds to g although g' does not speak of the derivability of sentences but instead merely of the properties of numbers. We now consider a sentence g' which is Goedelian relative to J_2, i.e. to the system of premises available to C at t_2. If J_2 is actually given, then, by using Goedel's method, such a sentence can be actually constructed. (But even if, at t_1, we do not know, for some reason, the state of C's information at t_2, we do know that a Goedelian sentence relative to this state must exist, and somebody may correctly guess it, by some stroke of luck, before t_1, and may incorporate it into q_1.) Now, as Goedel has shown, the true statement h which asserts that a Goedelian sentence such as g' is undecidable in a system J_2 is, in its turn, also undecidable in J_2. But a sentence which is undecidable in J_2, must be undecidable in every sub-system of J_2, i.e. in every weaker system. Now we shall clearly assume that J_1 must be a sub-system of J_2; for this is a consequence of the assumption that C has a cumulative memory so that it has at t_2 at least its earlier information at its disposal (which we assumed to be true). But then, h cannot be decidable in J_1. Now our near-Goedelian question q_1 is equivalent to h; accordingly, q_1 is undecidable in J_1, and C cannot answer it, except by signalling ' I do not know yet.'

---

[1] See note 2 on p. 180, above.

I shall now describe more fully a conceivable mode of operation of C. This description is intended as an illustration rather than as a list of conditions to be imposed upon C; and it is chosen, partly in such a way as to give C considerable powers of ' learning ', partly in order to avoid unnecessary complications. As far as I can see, there is no reason to believe that we have to impose any conditions at all upon C apart from its capability of making number-theoretical calculations (in view of Goedel's assumptions) and its capability of deciding certain questions by ' yes ' or ' no ' according to whether they or their negations follow from a certain system of premises.[1]

We may perhaps conceive of C as possessing two kinds of openings —one for the supply of information, and one (or more) for the supply of questions, both in standardised form. They may be inserted, for example, in the form of numerical expressions (one at a time) since statements can be correlated to numbers[2] by various methods, e.g. by ordering them in some lexicographical manner. We may assume that, after inserting a question a shutter closes the opening for information, as well as the one into which the question was inserted, until the question is decided; if no decision has been reached after a certain number of hours has elapsed, then, we may assume, the opening for questions is opened again, so that a new question may be inserted; and if this is done—but not otherwise—then the work on the old one stops. (If no new question is inserted, the work on the old one continues indefinitely or until a solution is reached.) Whenever the question (or, if more than one, all the questions) on which the machine is working has (or have) been answered, the information shutter opens, and we can, if we wish, supply new information before asking a new question. But as long as *C* works on a question, no new information can be inserted.[1] The actual calculation process may be conceived in this way : part of the machine continuously works out one demonstration after the other on the basis of the mathematical and logical principles built into it—for example, by proceeding from the shorter demonstrations to longer ones, and within the same length by some kind of lexicographical order. Another part arranges the so proved conclusions according to some order, and replaces, more especially, the proved conditionals as conditionals between two numbers (the numbers correlated to antecedent and consequent). A third part which works quite independently forms the conditionals whose antecedent is the number of the shortest sentence equivalent to the conjunction of the total information so far received, and whose consequent is the number of the question (or the questions) on which the machine is just working. And a fourth part, ultimately, compares these with the proved conditionals, and activates the ‘ yes ’ (or ‘ no ’) signals if it finds agreement. In this way the machine will, if given enough time, answer ultimately every answerable (derivable) question (and, indeed, by producing the shortest possible proof of it). It is clear that, if undecidable questions are inserted, the machine will work as long as it is permitted to.

---

[1] I think that it may be possible to build calculators capable of solving Goedelian questions—for example, by way of copying a method which operates not only with derivability but with truth, in Tarski's sense—and of adding the solution to its premises. Such a calculator may, whenever it is stimulated by a Goedelian question, ' grow ' in consequence of this stimulus into one whose system of premises is strengthened. But even such a calculator would, for example, be incapable of ' knowing ' at any moment whether its present system of premises was consistent—it could ' know ' such a thing only of its old premises after strengthening them.

[2] The significance of this fact was first emphasised by Goedel's and Tarski's ' method of arithmetisation '.

A machine constructed on lines such as these may continuously add to its own powers of answering questions speedily, even if it does not receive new information from outside. It is also clear that we can assume, if we wish, that the asking or not asking of a question q_i at t_i may make a difference to the speed with which the machine may later investigate a related question, owing to the fact that certain conditionals have been already formed (by what we have called the third part of the machine) and compared with certain conclusions (fourth part); in other words, we can assume that partial results obtained are stored for future use, and that they speed up all later calculations for which they can be used.

---

[1] New information can therefore be inserted only after a definite ‘ yes ’ or ‘ no ’ has been given to the last answer. Thus if the machine has been working on an insoluble question, and we wish to add new information, we must first stop its work on the insoluble question by inserting (after the period mentioned) some trivial question to which the machine will reply quickly. The point of this is (*a*) that the question always comes after the information, leading to a specific reaction of the machine (its work on the question), (*b*) that no other kind of interference (switching on or off, for example) need to be considered except the supplying of information or of questions.

If we now further endow the machine C with the means of storing, as information about itself, every question (answered or unanswered) which it has received, with a mark of the time of its receipt and of the time when it was answered etc., then we may make it possible for C to answer the near-Goedelian question q_1 as soon as such an answer has no longer any predictive value. For if q_2 was not inserted at t_2, the machine may 'find out this fact' on the basis of its storage of information, and therefore answer q_1, immediately after t_2, by 'yes' (since q_1 is now vacuously satisfied). And if q_2 has been asked at t_2, then the machine (provided it is capable of continuing to deal with one question—viz, q_2—even after another—viz, q_1—has been asked) might answer 'yes' to q_1 if, and immediately after, its work on q_2 was brought to a premature end (by insertion of a new question). Of course, both answers would be non-predictive.

These would be answers to a near-Goedelian question; but, of course, there is no contradiction in the assumption that C can answer even a Goedelian question after it has been added, in this or a similar way, to its information; for a Godelian question non-answerable in J_i may always be answerable in a strengthened system J_{i+1}.

So much about our near-Goedelian question q_1. We now wish to discuss a logically weaker although closely related question which C may be asked at t_1 instead of q_1—the semi-Goedelian question q_1'. It is the following question put to C at t_1: 'If you are asked, at t_2, the question q_2 (and allowed time until t_3 for an answer) will you or will you not before t_3 react with "yes" or "no" to q_2?' As before, we assume q_2 to be a Goedelian sentence relative to J_2.

Now the difference between the near-Goedelian question q_1 and our new or semi-Goedelian question q_1' lies in the introduction of the time-parameter t_3; and this is a very considerable difference. For as we have seen, it is impossible, according to Goedel, that h—or q_1—is derivable in J_2 (unless J_2 is inconsistent or at least false); h cannot, therefore, be among the information of which C is possessed at t_2 or at t_1. But our new question q_1 is logically weaker than h, and it is one whose answer may, without contradiction, be contained in J_1;

for example, J_1 may contain the (true) information that the question q_2 is too difficult to be solved before[1] a certain instant t_3. (In other words, C cannot, at t_1, 'know' that q_2 is insoluble in J_2; but it can 'know' that q_2 is, if at all, extremely hard to solve.) Such an information may therefore simply be given to C—say by a superior[2] machine C^+ which 'knows' that q_2 is undecidable in J_1, and therefore *a fortiori* that it cannot be derivable by C before t_3.

This consideration shows that we can no longer say of our new question q_1' that C cannot solve such a question, whatever its (true) information at t_1 may be.

Nevertheless we may say this: given the true information J_1—however full—we can always determine an instant t_3 so that our semi-Goedelian question q_1' is unanswerable for C on the basis of this information.

This follows very simply from the fact that J_1 cannot contain information strong enough to imply that q_2 is *too difficult however we may choose* t_3, for such an information amounts to saying that q_2 is insoluble, and this, we have seen, cannot be 'known' to C at t_1. Accordingly, even the strongest information which C may possess at t_1 can only be one which informs C—or which would permit C to deduce before t_2—that q_2 cannot be solved before a certain definite instant t_k. Given J_1, and other information about C, this instant t_k can be calculated. (But even if we are not in the possession of this information, t_k must exist and somebody might guess it, by some stroke of luck). And if we choose t_3 so that $t_k < t_3$, then q_1 must be unanswerable.[3]

---

[1] It may be remarked that, were J_1 to contain the information that no span of time shorter than a given time interval Δt is long enough for solving q_2, no matter at what instant this question is put, then this information would be contradictory, considering C's working methods as described above; for, owing to this, *every* answerable question may be answered as quickly as we wish if only it was asked sufficiently late in the 'life' of C or, for example, after C has just solved a closely related question.

[2] I am inclined to the conjecture (of which our considerations are completely independent) that C can obtain the information in question (that q_2 is too difficult to be solved before t_3, or something of this kind) only ready made from an outside source; that is to say, that it cannot deduce it from normal physical initial conditions (about its wheels etc.). The main reason for my conjecture is this. Since, in working on such a problem related to q_2, C would continually improve its powers for solving a closely related problem, it could never be sure, as it were, whether the solution of q_2 was not just 'round the corner'. See also the next note.

[3] In a system S_1, in which a Goedelian sentence g' (which asserts some property of *all* integers) is not derivable, its instances for every given integer n can nevertheless

As before, a semi-Goedelian question may well be answered in a non-predictive fashion, and even in three different ways; for an appropriately constructed *C* might answer ' yes ', as before, either if it finds after t_2 that q_2 has not been asked, or if it finds that its work on q_2 was brought to a premature end (which now means, before t_3); and in addition, it might answer ' no ' after t_3; for in this case it can ' find out ' after t_3 the historical fact that q_2 has not been answered before t_3.

If we now briefly compare the two questions—the near-Goedelian q_1 and the semi-Goedelian q_2—then we can say that the latter possesses to a higher degree than the former the character of a concrete question about a physical reaction of *C* while the former is more like a question about the general physical capacities of *C* or about *C*'s general physical state. Logically considered,[1] q_1 has greater similarities to a purely mathematical question than q_1'. It is for this reason that it may have been worth while, in spite of the trouble involved, to have discussed the more concrete semi-Goedelian question q_2' also.

Our general result may be summed up in this way. However full *C*'s information, there exist always questions about its future reactions to physical stimuli which *C* cannot answer—at least not before the events in question have become part of *C*'s historical past.

As before (towards the end of our discussion of the Tristram Shandy argument), we may raise a question about the relations between *C* and another computor C^+.

At first sight, one might get the impression that the situation is now different, and that we may put a question about *C* to C^+ which C^+ cannot answer, viz. the following question q_1^+: ' If *C* is asked at t_2 the question q_2, will it answer "yes" or "no" to q_2 ? ', with q_2 chosen in such a way that it is Goedelian relative to the information available to C^+ rather than to *C*. How, one might think, could C^+ answer

---

be derived. This means for a calculator *C* that it can derive, given enough time, the instances of g' for any *n*, however large; nevertheless there must be, at any given instant of time, an integer beyond which *C* has not proceeded in its derivations —or in other words, there will, at any instant of time, be a limit to the ' knowledge ' of *C* concerning *g* . This must also hold if we consider, instead of *C*'s growing knowledge of the properties of integers in relation to g', *C*'s growing knowledge of the degree of difficulty of solving a Goedelian problem (where this degree may be measured by an estimate of the instant of time before which no solution can be expected).

[1] Especially from the point of view of the Goedelian method, the difference between the two questions is immense—q_1 corresponds to a formally undecidable question, while q_1' does not.

such a question about C which involves a question it cannot solve itself? In this way we might be led to think that we have found a method of asking a machine C^+ questions about *another* machine which C^+ cannot answer, however fully it is informed about C. This, however, is not the case. C^+ can be informed of the fact that C is incapable of deciding q_2, without deriving from this information the kind of thing which it must not derive without becoming involved in contradictions, viz, that it itself—C^+—cannot decide this question q_2.

Thus we find, as before, that if C^+ and C are equipped with exactly the same information (and logical outfit) they will arrive at the same results, or will fail to arrive at these results in the same manner. But while the fullest information available to C still gives rise to questions about C which C cannot answer, it appears that C^+ can possess, in principle, such information about C (and especially about C's limitations) as to enable it to answer every physical question about C.

Since all this appears to hold for a C^+ which is not constructively superior to C, it clearly suggests that C's inferior knowledge is due to a logical impossibility of imparting to it information about itself without, by this very act, making this information obsolete. Indeed, if we inform C that the information which it now possesses is consistent, then one of two things may happen. The first possibility is that we mean by ' now ' the information possessed by C before we told it of its consistency. If ' now ' was meant in this way, then no difficulty arises; but C will still not ' know ', after we gave it this new piece of information, whether the total information which it now possesses is consistent.

If, on the other hand, we make it clear to C that ' now ' includes the piece of information about its own consistency—thereby trying to give C really up-to-date information about itself—then, according to Goedel, we attempt that which is logically impossible; for then the system J will imply its own consistency which, as Goedel has shown, can at once be used to derive a contradiction from J.

(III) By the name ' *Oedipus effect*,' I wish to allude to a certain aspect of the story of Oedipus whose fate was predicted by the oracle, with the result that the prediction induced Oedipus' father to those very actions which ultimately brought about the predicted events. The effect of a prediction (or a similar piece of information) upon the events or objects to which the prediction refers—for example, by promoting or by preventing the predicted events—I have called the 'Oedipus effect'.[1] One would expect the Oedipus effect to be a peculiarity of the social sciences and their subject matter, and that it is completely foreign to physics ; but this is not the case. As we have seen at the end of both our discussions of the Tristram Shandy argument and the Goedelian argument, the passing on to *C* of information about its own state is liable to interfere strongly with this state, and thereby to destroy the predictive value of the information.

This leads us to ask the question : Is it at all possible for *C* to obtain the information about its own past or present state which it would need in order to predict its own future state successfully ?

There is no difficulty in assuming that *C* has very full information about its immediate past ; for it may simply receive such information from another machine C^+ which is capable of measuring *C*'s state. The point is that the receipt of such information will change the state of *C*—and since *C* is a computor, constructed to work upon the information received, the interference must be a very strong one, not to be neglected. It therefore appears that we may be faced with an insurmountable difficulty in the way of all self-prediction. I believe that this is indeed the case.

The situation we are considering is very similar to the one we called the Tristram Shandy argument. But the objections raised against that argument no longer arise. The reason is this. The Tristram Shandy argument refers to the impossibility of formulating a complete self-prediction of a future state ; but it (*a*) must postulate restrictive assumptions about the calculator's power of formulation, and it (*b*) fails even then because it is at least dubious whether determinism demands a *complete* state prediction. The 'Oedipus effect', or rather our argument based on its utilisation, is concerned, on the other hand, with the impossibility of complete self-information ; and this is why it is here neither necessary (*a*) to introduce a special postulate concerning the formulation of this information (because a physical predictor furnished with physical theories can only operate on the basis of a purely physical initial description) ; nor (*b*) to doubt that the impossibility of completing the information invalidates at least some of the predictions based upon this information.

Thus we have merely to establish our thesis that *C*'s information concerning its own physical state must be incomplete at any instant of time, while another machine C^+ may possess complete information about C.

---

[1] Psycho-analysis is interested in another aspect of the same story, but it might profit from studying this effect—more especially, the influence of the analysis (the interpretation) upon the production of confirming material.

We have already seen that C's state is strongly interfered with if it is given any information by C^+. Still, this does not completely settle the problem; for C^+, one might think, may foresee this effect; and a very superior predictor C^+ might, instead of supplying C with accurate information about itself, work out a piece of information specially designed to induce C to predict a certain t_i state and, at the same time, designed to interfere with C in such a way that, at the instant t_i, C will in fact be in the state predicted by C.

I shall try to show that it is impossible for C^+ completely to succeed with this ingenious piece of trickery (which is so similar to what the Oracle achieved). No doubt, C^+ may, similar to the Oracle, induce C to do all kinds of things, by inducing in C certain states; and among the things which C^+ may achieve in this way may be correct self-predictions produced by C. But these self-predictions cannot be of the character of a complete self-information, and therefore not a valid basis for further correct self-prediction; accordingly, there will be always predictive questions about itself which C will be unable to answer correctly.

The simple reason is that C's self-information of its state at t_i, in order to be correct and up to date, would have to contain, as one of its parts, a physical description of that very self-information, using physical symbols of the same kind as in the rest of its information. But this, as we have seen before (see the ' lemma ' in our Tristram Shandy discussion) is impossible for any finite piece of information; and an infinite piece of information cannot, of course, be completed at any instant of time, so that it cannot become the basis of a prediction.

I consider this argument conclusive. It also answers the question whether C cannot, by way of some feed-back mechanism, possess some kind of immediate knowledge about its own state. Quite apart from the fact that a feed-back would inform C always after the event, it is clear that C can act, i.e. calculate, only on the basis of some formulated initial information, and that this must be always deficient for the reason given, if it is about C's own state.

In order to illustrate this point further, I shall discuss the objection that C may be constructed in such a way as to be able to use all incoming information twice over—once in the normal way, and once as a piece of information about C's having been interfered with at certain moments by the insertion of a tape of a certain shape.

Now it turns out, of course, that even this does not change things. If C can use its information *only* twice, then the self-prediction based upon the second use will be as inaccurate as one based upon such a second information received from C^+; for by acting upon this second information, C must act in a way it cannot foresee, having been informed by the second information only about the receipt of the first (but not of the second upon which it acts). If, in order to remedy this, we construct C in such a way that it can use also the second information twice, and so on *ad infinitum*, then C will never be able to begin predicting, owing to the incessant stream of self-produced new information.

(IV) To sum up our three arguments:

They all agree that calculators are incapable of answering all questions which may be put to them about their own future, however completely they are informed; and the reason for this appears to lie in the fact that it is impossible for a calculator to possess completely up-to-date initial information about itself. ('Know thyself', it seems, is a contradictory ideal; we cannot fully know our limitations—at least not those to our knowledge.)

For we have seen that a calculator C can be given, by an observing machine which may or may not form part of C, full information about C's past states but never about C's present or future states. And although the information about C's past states may be complete as far as C itself is concerned, it can *not* comprise sufficiently full information about C's environment—about some closed system which surrounds C and with which C interacts. For the machine which supplies information to C, and which thereby strongly interferes with C cannot in its turn possess full self-information and cannot, therefore, supply such information to C.

This holds for a calculator C. For complete prediction machines, our older argument holds (see section 7): a region in which one or more predictors exist cannot be exhaustively measured by any of them for predicting purposes owing to the strong interference of the measurement with some of the objects measured.

On the other hand, there does not appear any serious obstacle for a calculator to explain in any desired detail its own past states. (For explanation, in the scientific sense, is the same as prediction, except that the demand is dropped that the deduced statement must be obtained *before* the event which it describes.) In order to give such an explanation, a calculator capable of ‘ learning ’ may use its ‘ historical knowledge ’—its records of past stimuli and its reactions to them. We may conjecture that only a calculator capable of ‘ learning ’ in this sense can explain in detail its own past, and that it can do so only because its powers (its ‘ knowledge ’) at the time at which it gives an explanation exceeds its powers (its ‘ knowledge ’) at the time for which it gives an explanation.[1]

At any particular instant of time t_i, there will exist, for any machine C, a characteristic time-interval $\Delta\ t_i$, viz, the shortest interval between the instant t_{i-1} at which the latest event E has occurred which C can ‘ know ’ at t_i owing to its having become an event of the past, and the instant t_i, that is to say, the one at which C's ‘ knowledge ’ of E is forthcoming in the shape of an explanation or description. We may call this characteristic interval the ‘ specious present ’ of C. (In other words, the ‘ specious present ’ of C is the minimum time it takes C to ‘ know ’ what has happened to it—not the minimum time it takes C to react to a stimulus ; this, in general, will be much shorter.) The length of this interval will depend on C's general construction, but also perhaps on acquired ‘ learning ’—the use of its past results may, for example, help to shorten it. (It may also lengthen it, by overburdening C's ‘ memory ’.)

The ‘ specious present ’ of C divides *time* for C into past and future —not for C's ‘ consciousness ’, for this it has not, but for C's functioning. Time, for C's functioning, is divided into that part for which every question asked can, in principle (i.e. given sufficient factual information) be answered, and that part for which there are questions whose answers cannot, in principle, be ‘ known ’ by C. In other words, C's time is divided by its ' specious present' into a closed or fixed past, and an open not fully determinate future. (And this does not only refer to C's knowledge of itself, but also to its knowledge about its ' closer environment'.)

---

[1] Professor F. A. Hayek suggested to me several years ago that no calculator can *explain in detail* its own way of functioning ; his argument was that explanation in detail can only be achieved by a system of a ‘ higher degree of complexity ’ than the system to be explained. My results may perhaps be interpreted as agreeing with this argument ; but they seem to show that the argument does not apply to all calculators and that we must distinguish between self-explanation, which in principle is possible, and self-prediction, which is not.—Professor E. Tranekjaer-Rasmussen (see his *Psychologie und Gesetzerkenntnis*, in *Gesetz und Wirklichkeit*, ed. by S. Moser, Innsbruck, 1949 ; see esp. p. 121) developed a very interesting argument purporting to show the impossibility of a physiological explanation of the experiences of understanding and explaining. I believe, however, that this argument is again incapable of showing more than that a self-explanation of *present or future* states is impossible ; it does not, as far as I can see, conclusively show the impossibility of a physiological explanation of other peoples' past and future, or of one's own past, experiences of the kind mentioned.

9

If our argument is correct, then we have two different results. One is logical. It asserts that even if any prediction task given is capable of being carried out by some predictor, there will be no predictor capable of carrying out every task, since no predictor can fully predict its own closer environment. It further asserts that the deterministic assumption that an adequate predictor can be constructed for every prediction task leads to the conclusion that an infinite series of predictors with strongly growing powers (i.e. a series which is non-convergent) can be constructed. The other result refers to certain systems of mechanics and asserts that they are incompatible with this conclusion. It is a consequence of the latter result that we can assert that classical mechanics is not deterministic, but must admit the existence of unpredictable events.

If this is correct, then Laplace's determinism, and that of others who were influenced by the *prima facie* deterministic character of classical mechanics, is based upon a misinterpretation. I suggest that this misinterpretation is due to the tendency of attributing to Science (with a capital S) a kind of omniscience ; and that this theological view of science ought to be replaced by a more humanistic view, by the realisation that science is the work of ordinary humans, groping their way in the dark. In doing so, we may sometimes find something interesting ; we may be astonishingly successful ; but we shall never get anything like ' the whole truth '. Our theories are not descriptions of nature, but only of some little feathers which we plucked out of nature's garb, more or less accidentally.

One warning may be especially needful in this connection, owing to the fact that the present paper may be otherwise misinterpreted as an expression of my agreement with the now often laboured analogy between a human brain and an electric calculator.

In so far as man is a predictor, my modest results are, I believe, applicable to man and the human society.

But I certainly do not intend to suggest that the parallelism carries far. I do think that man is not only a predictor, but more. I think that even in so far as our purely intellectual activities are concerned, we have hopes, fears, interests, and problems. We are not calculators, or in so far as we are calculators, we are miserably bad ones. Every ordinary adding machine is superior to most of us. Indeed, we would not construct multiplication tables, and systems of arithmetics, if our brains could calculate. We construct methods outside our skins to calculate with pencil and paper ; and we construct electronic brains because we have no such brains ourselves.

Thus we are not calculators. But we are constructors of calculators. We make them because we are interested in problems whose solutions are beyond our limited calculating powers, and because we are fascinated by the new problems which the construction of calculators presents to us. Our fundamental intellectual impulse is that of searching for difficulties ; it is even one of inventing difficulties, in order to overcome them.

A calculator may be able, for example, to produce proofs of mathematical theorems. It may distinguish theorems from non-theorems, true statements from false statements. But it will not distinguish ingenious proofs and interesting theorems from dull and uninteresting ones. It will thus ' know ' too much—far too much that is without any interest. The knowledge of a calculator, however systematic, is like a sea of truisms in which a few particles of gold may be suspended. It is only our human brain which may lend significance to the calculators' senseless powers of producing truths.

To put this argument in a more formal way, it is the function of every theory to divide all statements pertaining to the subject matter in question into three distinct classes—those which the theory asserts to be true, those which it asserts to be false, and those about which it does not make any assertion. It is for this reason that an inconsistent theory is useless ; for it does not achieve such a division, but asserts every statement (and therefore also the negation of every statement). An inconsistent theory is useless because it asserts too much.

Now a good (i.e. a consistent) calculator is certainly not useless, but able to achieve such a classification. However, it still asserts too much ; for example, with every moderately useful statement such as ' $2 + 1 = 3$ ', it will also contain an infinite series of statements ' $2 + 1 \neq 4$ ', ' $2 + 1 \neq 5$ ' . . . and another infinite series of statements such as ' $2 + 1 \neq 3 + 1$ ', ' $2 + 1 \neq 4 + 1$ '. . . And if it is made by some method or other, to derive automatically all the consequences of a theory one after another, then it will still possess no method of picking out the interesting or important ones, nor a method ensuring that it produces even one of these within any definite interval of time. In the infinite series of the statements in their order of production, the probability of hitting upon one which is interesting (by any reasonable standard) will be zero.[1]

It is only the human brain which can create interests, purposes, problems and ends—even within the comparatively narrow field of its intellectual activities.

10

The general upshot of our considerations appears to me to be the restitution of the naïve view of the world, described as the 'common-sense view' at the beginning of section 2—the view that there are events which can be predicted, or which are 'determined', and other events which cannot be predicted and are not 'determined'.

But our considerations even suggest something like a reconciliation between this view with the other—the 'more sophisticated' view that it is as a rule only lack of knowledge which makes us believe that events are unpredictable.

This reconciliation can be brought about if we realise that it is the existence of knowledge in the physical world which creates the kind of indeterminism we have been discussing. Our knowledge may conquer new problems of determination. But in doing so, it will create new predictive problems which it cannot solve. For it cannot foresee its own future conquests.

KARL R. POPPER

---

[1] Another formal argument would be this. We learn by mistakes ; and this means that when we arrive at inconsistencies we turn back and reframe our assumptions. And in applying this method we go so far as to re-examine assumptions even of a logical nature, if necessary. (This happened in the case of the logical paradoxes.) It is hardly conceivable that a machine could do the same. If its creators were uncautious enough to equip it with inconsistencies, then it will simply derive, in time, every statement that it can form (and its negation). We may perhaps equip it with a gadget which warns it, in case it derives '$0 = 1$', and makes it abandon some of its assumptions. But we shall hardly be able to construct a machine which can criticise and re-adjust its own methods of derivation.

SCIENCE AND PHILOSOPHY: IMPLICATIONS OR PRESUPPOSITIONS?

I

Possibly the commonest notion among scientists about the relations of science to philosophy is epitomised in that phrase so often heard, ' the philosophical implications of science '. It is often supposed that science stands as a body of knowledge independent and self-justified, and further that some scientific conclusions are relevant to philosophical problems. Against this I wish to outline the view that science has no philosophical implications (in the strict sense); but that conversely, when considered as a body of knowledge about nature, it requires certain philosophical presuppositions, and therefore depends on philosophy for certain basic assumptions which it takes for granted.

The common presumptions that need examination are (*a*) that the method of science is self-evidently valid and makes no assumptions whose justification lies beyond the scope of science, and (*b*) that science and philosophy not only deal with the same subject-matter but treat it from the same point of view. The key to the problem is, I think, the analysis of scientific method. For this method must depend on the subject-matter of science and on the point of view from which it is approached; and the method determines the kind of knowledge that can finally be attained. Thus by considering the method and working backwards, as it were, we can arrive at the scientific point of view; and by working forwards we can assess the types of conclusions that science can reach, their reliability, their presuppositions, and their status as descriptions of reality. To discuss the relation of science to philosophy by starting with the conclusions of science, rather than its method, is to build on sand; the first task is to determine exactly the status of those conclusions as a form of knowledge, that is, to determine precisely what they mean, how they are related to the evidence, and to what questions the evidence is relevant, i.e. what is the point of view of the investigation. We ought to begin, then, with a study of the method and presuppositions of science, not with its conclusions. In what follows I shall confine my attention to physical science, since its use of measurement and mathematical reasoning makes its method especially simple to describe and assess. Biological and psychological sciences would need modified treatments, but I do not think that the general results would be very different.

2

For our present purposes, the method of physical science may be briefly characterised as follows : (i) It is concerned with observations on material systems. These observations are measurements ; that is, physics confines itself to the quantitative aspect of phenomena. It is concerned with the description and correlation of phenomena. A physical experiment or set of observations commonly yields two parallel columns of figures, representing corresponding values of two variables—such as the pressure and volume of a gas, or the position of a star at various times. (ii) The next step is to correlate these figures, to find some mathematical expression with which each and every pair of values is in approximate agreement. If such an expression can be found, it is taken as at least an approximation to a general law describing the concomitant variation of the two variables. Implicit in this step is a peculiar kind of generalisation ; as we shall see, the analysis of the validity of this generalisation uncovers the chief presupposition of the inductive method, which is the main point of contact between science and metaphysics. (iii) The empirical laws thus obtained for various phenomena are unified by theories. These consist fundamentally of equations. The equations are such that one can derive from them, purely by mathematical deduction, expressions that agree within experimental error with the empirical laws. Thus (to take an early example) the theory of universal gravitation consists of an equation (the inverse-square expression) from which may be deduced equations that agree with the orbits of the planets round the sun, the orbit of the moon round the earth, and the motion of falling bodies on the earth. A theory is taken to be supported by the evidence in so far as deductions from it agree with the laws derived from observation. (iv) The theoretical interpretation is often, but not necessarily, accompanied by a model. Thus the theory that accounts, *inter alia*, for the relation between the pressure and volume of a gas is expressed in terms of a model of the gas as a swarm of molecules in motion ; and the equations that account for the diffraction, interference and refraction of light are commonly attached to a picture of light as a train of waves. Much of the ordinary physical picture consists of such models—electrons, photons, waves and so on—and it is of great importance to decide whether such models are necessary and fundamental to physics, and what exactly is their status. We have to enquire in what sense exactly we are to take statements like 'hydrogen gas consists of molecules, and each molecule consists of two protons and two electrons, and may emit photons if suitably treated'. This is the second point of contact with philosophy; for it is statements like this that are taken as the conclusions of science and sometimes supposed to be of importance to philosophy, and they can only be assessed by a logical investigation of the reasoning that leads to them.

This brief sketch of the essentials of physical method may serve to bring out what is needed from the present point of view, which is concerned with the logical status of physical laws and theories. Our treatment is not intended as an historical account, which would have to describe the dialectic of science, considering the intertwining of experiment and theory, and of theory and model. Still less is it a psychological account, which would be concerned with such topics as the role of imagination in the construction of theories. The point of view of our present approach is to determine the reliability of scientific conclusions, their relation to the evidence adduced for them, and their dependence (if any) on extra-scientific presuppositions. We are concerned with scientific method in the sense of a logic of proof, not in the sense of a method of investigation.

The method of physics is admirably coherent. We begin with measurements on material systems; since these give us sets of numbers as our primary data, the empirical laws that we derive from them naturally take the form of functional relations, or equations; and equally naturally the theories, since the laws must be deducible from them, consist of equations. The method indicates at once the point of view of the investigation and the type of conclusions that emerge. The point of view, clearly, is the observation and correlation of the quantitative aspects of phenomena. The conclusions, which comprise physical knowledge, are mathematical equations, from which (directly or indirectly) the laws of concomitant variation of the measurable aspects of phenomena may be deduced.

We can now turn to the relation of physics to some of the branches of philosophy, expanding our account of physical method as it becomes necessary. We cannot generalise about the relation of physics to 'philosophy'; we must consider separately the branches of philosophy. These are, I take it, logic, epistemology or theory of knowledge, metaphysics, ethics, and (perhaps) aesthetics. Here I shall leave aside the last two, and concerning the first I shall only say that formal logic, which lays down the methods of valid inference from given premisses but has nothing to say about establishing the truth of the premisses, is for science simply a tool, as it is for all reasoning. Epistemology and metaphysics will detain us longer.

3

Epistemology, or theory of knowledge, is concerned with the relation of our cognitive experience to reality. It is distinct from the psychological enquiry into *how* we come to know, believe, assume, or invent what we call our knowledge—the mechanism, so to say, of cognition. Scientific method may contribute to the psychological enquiry, but not the epistemological, because it does not use the same point of view. It is concerned with the correlation of phenomena, not with the relation of the phenomena to reality.

Science can be regarded in two ways. It can be regarded, if we so wish, purely as a systematisation of sensory experience—a unified scheme of observations, laws, and theories ; the essential point being the correlation of phenomena and the pattern that results. In this sense, science is autonomous and needs no epistemology, as Whitehead has argued,[1] and a scientist who wishes to avoid epistemological enquiry can proceed with his science if he confines himself to this outlook. But it is a limited outlook, because the question of the *truth* of science, its relation to nature (the reality which is its subject-matter), has not been considered.

This leads to the second way of regarding science : as an approach to truth about nature. As soon as the question of the truth of science is envisaged, we are concerned with epistemology. The first question is, can science provide its own epistemology ? Clearly not, because its method adopts a point of view that does not lead to the right sort of conclusions ; scientific method is ideal for correlating phenomena, but useless for investigating their relation to reality. This will become clearer as we proceed.

Next, does science require a particular epistemology ? It seems, on the contrary, that science can be interpreted on any epistemology, and that the choice between them must be made on other grounds. Let us consider a few examples. For those who take a realist view in epistemology, our sensations give knowledge (using the word in the broad sense) of realities ; and however inexact and approximate it may be, this is genuine knowledge, not wholly unrelated to the nature of the things concerned. The type of knowledge we obtain on a given occasion naturally depends on the point of view we adopt. Suppose we are examining the conduct of a physical experiment, for instance. Then if we take up the physical point of view, we shall arrive at certain correlations of measurable variables ; but if, instead, we adopt the physiological point of view, we shall examine the experimenter and note the co-ordination of his muscles, respiration, secretions, and so on ; again, from the ethical angle, we shall be interested in the experimenter's motives in doing the experiment ; and from the metaphysical point of view, we shall consider him also as the efficient cause of certain changes. On the realist view, then, scientific knowledge appears as knowledge of material things *from a particular point of view*, and as an approach to truth within these limits. Its laws and theories correspond with reality, though to what extent is a further question.

---

[1] A. N. Whitehead, *The Concept of Nature*, Cambridge, 1920 ; *An Enquiry Concerning the Principles of Natural Knowledge*, Cambridge, 1919.

A realist epistemology, then, holds that the cognitions that we derive, ultimately, from sense-experience depend for their main features on the nature of whatever gives rise to the sensations. A quite different account is given in Kant's epistemology, according to which our cognitions are not of reality, in the sense of things-in-themselves or noumena, at all ; phenomena are quite different from noumena. This difference is partly due to the imposition of the spatial and temporal characteristics that we meet in all our sensations on the raw material of sensible experience (which we never experience in its rawness, so to speak). And it is partly due to the further imposition of the characteristics of thinghood (substance) and regular sequence of events, which give order to our experience and save it from being a mere flux of sensations. In other words, when we pick out from experience certain continuant things and certain laws of their behaviour, it is not because things-in-themselves or noumena are anything like that, but because our minds and senses work in such a way that we cannot help it. I am not here discussing the truth of the theory ; I am only pointing out that this theory of knowledge is equally able to find room for science. For, by this account, science is no more concerned with nature-in-itself than is any other human cognition ; in the scientific account, as in all human knowledge, the raw material has had the categories imposed on it ; it is a coherent systematisation of sense-data which all can agree upon, and that is all. Given the epistemology, there is no difficulty in giving an account of science.

Both these epistemologies recognise a certain order in our experience, and they treat science too as an ordered account of certain aspects of experience, as it certainly appears to be. A purely empiricist or positivist epistemology, however, attempts to reduce experience to a succession of sensations, making them the sole knowable reality. On a thorough-going empiricist view, the laws and theories of physical science do not reflect an ordered reality ; they are simply a superstructure by means of which we are able to predict future events with some degree of confidence from our experience of past events. The interest of science does not lie in general laws and theories, but in this tenuous connection of a future event with past events. Observations are to some extent unified by science, but the equations which achieve the unification are pure fictions. Again we see how an epistemology, developed without calling on the conclusions of science, can lay down its own interpretation of scientific investigation.

It appears, then, that the existence of natural science can be interpreted in terms of whatever theory of knowledge one has (on other grounds) adopted, and does not demand any particular theory. Science has in fact no epistemological preferences. It does not favour one theory of knowledge rather than another.

In this branch of philosophy at least, then, the conclusions of science have no bearing on those of philosophy. This is perhaps hardly surprising when one remembers the different points of view and different objectives of the two investigations. But it is worth stressing, because of the difficulty that some scientists have in seeing that besides the scientific point of view there are others, and that scientific method is not the only rational method. We have also to contend with the confusion brought into such discussions by those who equate the theory of knowledge with the description of scientific method. It cannot be said that scientists have been adequately helped by professional philosophers in this matter. So long as the paradox continues that philosophers give a clear and exact account of the general principles of scientific method, but no account at all of philosophical method, scientists and others will be disposed to imagine that the scientific method is the most reliable in all investigations and confirmed in their notion that philosophy is only out-of-date science. An examination of philosophical method that would show exactly where and why it differs from the inductive method would be a great advance.

4

In what follows it is necessary to make a choice of epistemology. My choice is the realist type of view—the kind of theory which asserts that we can reach some knowledge of reality, though with difficulty. It would take too long to argue this in detail ; here I will only say that the empiricist view runs into great difficulties over universals or ' abstract ideas ', and the Kantian view does not seem to do justice to the objectivity and immediacy of knowledge, while the realist view can surmount these difficulties ; the rest of the argument must here be left to the philosophers.[1] Fortunately, however, most scientists seem to be implicit realists, and will probably agree with this point of view. We shall, then, attempt a realist analysis of scientific method, and shall regard science not merely as a systematisation of observations, but as a description of nature (from its special point of view), claiming as such to be an approach to truth about nature.

Let us now refer to the account of physical method given previously, and in particular to step (ii). This step is the passage from the experimental data—commonly consisting of two parallel columns of figures, representing two concomitant variables, such as the pressure and volume of a gas—to a general law (Boyle's law, in this example) which is taken to connect the two variables on all similar occasions. Or we may think of the transition from a few points plotted on graph paper, to the continuous curve that we draw through them and regard as universally valid (for the given conditions). What is the logical status of this procedure ? We generalise from a few observations to a general equation ; we assert a general law on the basis of a few instances of it. We do this almost automatically, without considering the extrapolations involved—without adverting to the innumerable specimens of gas that we have not examined but which we assert will obey the same law.

---

[1] Cf. e.g. H. A. Prichard, *Kant's Theory of Knowledge*, Oxford, 1909, esp. chap. vi ; A. C. Ewing, *Idealism*, London, 1934 ; M. C. D'Arcy, *The Nature of Belief*, London, 1931 ; D. J. B. Hawkins, *The Criticism of Experience*, London, 1945.

It is an interesting kind of generalisation. It is not simply a summary of the evidence, such as we might produce if we were to measure all the policemen in London and conclude that they were all between five and seven feet high ; it goes far beyond that. It is not simply a shorthand version of the observations, for it is asserted as a general law applying to phenomena that we have not investigated. Again it is not like a statistical statement, such as we might make in reporting the proportion of deaths due to road accidents in a given year ; for this again depends on complete enumeration, and does not go beyond the evidence. Our inductive generalisation, by contrast, states more than is implied in strict logic by the observational evidence ; for by no process of formal logic can one deduce a general law, of the type $x = f(y)$, from a finite number of propositions of the type $x_1 = f(y_1)$, $x_2 = f(y_2)$, etc. The law is, from a logical standpoint, a *construction* from the observations ; or again it is an *interpretation*, since evidently the observations are treated not as bare data but as signs of some general law, which we discover by interpreting them. On what grounds, then, do we justify this method of generalising by construction or interpretation ? This is the fundamental problem of induction.

What we need is some principle which when combined with the observations will confer some likelihood on the generalisation based on them, and which we have independent grounds for believing to be true. The minimum principle required was worked out by J. M. Keynes in his *Treatise on Probability*, and later work has mostly taken its cue from him. Without delving into logical details, it may be said, translating Keynes' principle into terms of a realist philosophy, that the logically necessary principle is that *there is order in nature* ; that the behaviour of nature is not entirely capricious, or (what is the same thing) that some generalisations hold in nature. If we have antecedent grounds for believing that there is *some* order, we can use our scanty experimental data to determine *which* order ; if we expect some law to hold, we can treat the observations as signs pointing to that law ; and our generalising has a rational basis.[1]

This then is the kind of principle that the method of induction tacitly appeals to, and if induction is to be valid we must have grounds for believing it, or something like it, to be true. The next question is, what sort of grounds can these be? They cannot be purely logical; as we have noted, formal logic is concerned with valid inference from given premisses, not with the truth of the premisses. They cannot be derived from science, for all scientific conclusions rest on induction and so the argument would be circular. If there is any basis for induction, it must be derived from outside science—it must indeed be metaphysical.[1] If we believe that there is order in nature, it is on metaphysical grounds; in science the belief is presupposed, and we cannot therefore appeal to science to support it. No doubt most of us believe it by virtue of the unformulated metaphysic of common sense; but when we come to an exact analysis of the relation of scientific conclusions to the evidence adduced for them, the metaphysical origin of one of the great presuppositions of scientific method is forced into the light. We shall return to the question of its credibility in a moment.

---

[1] The use I am making of the principle of order in nature is not the same as J. S. Mill's use of his principle of the 'uniformity of nature'. Mill thought that the principle of uniformity would provide a sufficient condition of a generalisation to be certain. The present treatment introduces some such principle as that of order in nature merely as a necessary condition for an empirical law to have any likelihood of being true.

The same presupposition appears in the construction and testing of theories (step iii). A theory is presumed to be supported by the facts when deductions from it agree with empirical laws. But the theoretical equation is not, conversely, deducible from the empirical laws alone; it has to be constructed, not deduced. The argument is not of the following type, which is deductively correct: 'The equations p_1, p_2, p_3, etc. are empirical laws supported by observations; these equations together imply the equation q (a unifying theory); therefore q is supported by the observational evidence'. (I use 'implies' as the converse of 'is deducible from' throughout; if p implies q, q is deducible from p.) Theories are related to experience by a sort of inverted form of this argument: 'q implies p_1, p_2, etc.; but p_1, p_2, etc., agree with empirical laws based on observation; therefore q is supported by the evidence'. The conclusion here does not follow in strict logic from the premisses. Thus the validity of a theory is not deducible from the empirical laws that are taken to support it. Theories are in fact not deductions, but interpretations, as indeed they are commonly called. They are believed because we regard empirical laws as *signs* of a more fundamental order in nature, represented by the unified theoretical scheme which interprets them. This order must be presupposed. Thus the justification for theorising requires again the assumption of order in nature. It raises therefore the same metaphysical question as does the formulation of empirical laws.

---

[1] By adopting a realist epistemology we have ruled out the Kantian solution, already referred to, which would provide an epistemological basis rather than a metaphysical one. On either view, induction depends upon a principle that can be investigated only by philosophy and not by science.

At this point the work of critical philosophy, the logician's analysis, comes to an end ; the next stage consists in finding grounds for the presupposition of order in nature, and requires a constructive metaphysic. This part of the problem has not received adequate attention from philosophers ; metaphysics is unfashionable among them and they have generally stopped short after the critical analysis of induction. The way to look for a solution, it would seem, would be to study anew the great metaphysical tradition of the west—of Plato, Aristotle, Aquinas, Scotus and so on down to Whitehead. But this is not here our concern ; it is enough for our present purpose to conclude that, on a realist analysis, natural science (considered as an account of nature) uses a method that requires an extra-scientific presupposition, whose truth it is the business of metaphysics to examine.

It may be asked how there can be two different branches of thought —science and metaphysics—both dealing with nature as their object. The answer is that their points of view are different, and so also in consequence are their methods and the kind of conclusions they lead to ; just as in the study of human history we get different (but not necessarily incompatible) results when we adopt the points of view of economics, politics, religion, and so on. The point of view of physics is the correlation of measurable phenomena ; its method is inductive ; its conclusions are mathematically expressed laws and theories. The point of view of metaphysics (to put briefly the results of a long argument [1]) is directed to the beings that we know through phenomena, and the conditions for their existence and behaviour.; its method is not inductive but reflective ; its conclusions are the ultimate conditions for the intelligibility of experience. The two disciplines do not conflict because they do not deal with the same topics ; though both study nature, they do so from different points of view, asking different questions and reaching conclusions of different types. In particular, metaphysics is able to consider efficient causes and final causes, which are excluded from physics by its self-restriction to the measurable aspect of things. Thus metaphysics can give some account of the dependence of nature on its first cause ; it may well prove important to consider this in formulating metaphysical grounds for our belief in the order of nature.

---

[1] Cf. e.g. the author's *The Power and Limits of Science,* London, 1949, chap. vii.

5

Men of science have sometimes believed that the point of contact of physics with metaphysics was to be sought not in the root presupposition of physical method, as I have argued, but in the account of nature given by physics ; it has sometimes been supposed that the physical scheme of ultimate particles and general relations has in itself some metaphysical significance. This seems to me to be mistaken, because the physical account is derived from a different point of view from that of metaphysics. In the past, when mechanical models were still in fashion, physical explanations were sometimes expressed in terms that looked metaphysical, because the interactions of the physical particles, waves, and so on were metaphorically expressed in terms of efficient causation. A more exact analysis of the status of such physical models is helpful ; and the key to it is again the examination of the inductive method. We turn, then, to step (iv) of the physical method sketched above.

The status of physical models is much clearer today than it was fifty years ago. First of all, it has been realised, since the rise of quantum mechanics, that a tenable theory need not be associated with an imaginable model ; in fact physical models are not even possible in fundamental theory. As an illustration, consider the variable ψ of the wave-equation ; it corresponds to nothing that can be visualised, and is related to phenomena not by a model but by mathematical deduction. We still speak of particles, but they are not particles in the classical sense, because, as Heisenberg's uncertainty principle has shown us, they cannot be assigned a definite position, shape, or velocity. Physical theory is in fact fundamentally a system of equations, and the introduction of models is logically justifiable only in as much as they embody the correct equation. It is the equations that 'save the appearances' which are fundamental, from the point of view of the logical supporting of conclusions by evidence. Models, then, are not necessary to the logic of physics, however useful they may be to our imaginations in physical investigation. We must investigate a little further their relation to the mathematical scheme of theoretical physics.

A useful pointer is provided by the fact that widely different mechanical models are sometimes found to be associated with the same equation. An example is cited by Sir Edmund Whittaker :[1] ‘It happens very often that different physical systems are represented by identical mathematical descriptions. For example, the vibrations of a membrane which has the shape of an ellipse can be calculated by means of a differential equation known as Mathieu's equation : but this same equation is also arrived at when we study the dynamics of a circus performer, who holds an assistant balanced on a pole while he himself stands on a spherical ball rolling on the ground. If now we imagine an observer who discovers that the future course of a certain phenomenon can be predicted by Mathieu's equation, but who is unable for some reason to perceive the system which generates the phenomenon, then evidently he would be unable to tell whether the system in question is an elliptic membrane or a variety artiste.' This lack of a unique relation between model and equation suggests that the model is not necessarily a simple replica of the real system whose behaviour it simulates.

This impression is confirmed when we find that the behaviour of a given system may require different models according to circumstances. The fact that a beam of light may be treated by a particle-model in one experiment and a wave-model in another, according to the system with which it is interacting, indicates that neither model is a direct replica of the light beam. It suggests that only some, not all, of the characteristics of the models are the same as or similar to those of the reality. The light-beam has some characteristics in common with a wave travelling down a stretched string, and other characteristics in common with a travelling projectile, but does not share all its characteristics with either. In other words, the models are *analogues* of the real system ; for two things are analogous (in modern logical terminology[2]) if they have some, but not all, characteristics in common. They are not replicas, but analogues.

This view of models as providing analogies is confirmed if we reflect on their logical status, revealed by the way in which they are related to theories and so to the observational evidence. A successful model for a given physical system is one that leads to equations identical with the empirical laws or theoretical expressions, derived from observation in the ways already outlined. But this agreement does not imply that the system is exactly like the model ; only that it is like it in some respects. And this is the definition of an analogue. This conclusion might indeed have been reached on methodological grounds alone, without appeal to the physical experience summarised above.

---

[1] *The Beginning and End of the World* (Riddell lectures, 1942), Oxford, 1942, p. 17.

[2] Unfortunately the usage is different in metaphysics : cf. E. L. Mascall, *Existence and Analogy*, London, 1949, etc.

Models, then, are best regarded as analogies ; they are to be taken neither as exact copies of reality on the one hand, nor as sheer fictions on the other, but as analogies. This conclusion is extremely useful in dealing with a variety of pseudo-problems. One such was produced by the misunderstanding of the situation when it was found that the behaviour of light required two analogies, the wave and the particle, according to the type of experiment performed. It was supposed by some that science had led to two incompatible views about the nature of light, and naturally this caused considerable perplexity ; however, the puzzle vanishes if we remember that the wave and particle models need not be taken as exact replicas of a beam of light, but as analogies for its behaviour, and that the use of different analogies for its behaviour in different circumstances is quite legitimate. The ' luminiferous ether ', similarly, is an analogy, and it is no longer puzzling that it has some of the properties of a material medium (inasmuch as it ' transmits waves '), but not all of them, as was shown by the Michelson-Morley experiment. Indeed it may well be doubted whether the Michelson-Morley experiment would have assumed crucial importance if the status of theoretical models in physics had been adequately studied at the time ; the conclusions laboriously reached by physics might have been speeded by more attention to methodology. Incidentally, the superseding of the ether in general relativity by the curvature of space-time, which fulfils the same functions, is a good example of the way in which a mathematical theory may dispense with models.

Finally, from this view of models we can learn how to answer such questions as ' Do atoms exist ' ? The correct answer appears to be : ' We don't know ; but at any rate something analogous to the modern model of an atom exists. We don't know how close the analogy is, either ; but at any rate it is closer than the analogy with Dalton's model of an atom, and that was closer than the analogy with Democritus' atom.' At a time when our knowledge of ultimate particles is changing and when exciting news may be expected, it is well to be clear as to what exactly we mean when we say that a piece of iron is made up of atoms and so forth.

6

It is tempting to enquire whether the clarification of the relations between science and philosophy that is now possible has any significance for the study of the history of science. Perhaps one who has no claim to professional historical knowledge may be pardoned if he wonders whether, in the study of the history of science and its place in human thought, enough attention has been given to the fact that science was practised very successfully before its method or its relation to metaphysics was properly understood. To take one instance only : when Newton concluded that on his mechanical principles the solar system would be unstable, he supposed that matters would be set right by the direct intervention of God ; when Laplace showed that on Newton's principles the solar system would in fact be stable, it was apparently supposed by some people that there was no longer any need for ' that hypothesis ', the First Cause. But science and metaphysics work from different points of view ; science can never by itself conclude that God exists, because it is concerned only with phenomena ; while metaphysics concludes that there must be a first cause in order to account for the bare existence of a solar system, and for its following any laws at all, whatever those laws may be. Laplace's followers were wrong if they admitted only explanations in terms of laws ; Newton was right in admitting also explanations by causes, but wrong in importing them into scientific description. This is only one example of the confusion between laws and causes that seems to have prevailed in the minds of scientists until relatively recently ; physical method was not clearly distinguished from the more familiar analysis of experience in terms of causal activity of persons and other agents acting for definable ends.

Another example of confused thinking over scientific method is the common notion that the general acceptance, in the course of the seventeenth century, of the Copernican system, as against the Ptolemaic, was a triumph of scientific thought. It does not appear to be widely realised that the view that, taking the fixed stars as the frame of reference, the sun is stationary and the earth revolves round it, was supported by no evidence that would be accepted as conclusive by modern scientific standards. There were at the time no direct observations of the motion of the earth relative to the fixed stars ; nor did the Copernican model give markedly better agreement with observation than the Ptolemaic. The Copernican argument was in fact based largely on the greater *simplicity* of the model, compared with the complexities of the Ptolemaic system. This appeal to simplicity was not in fact based on scientific grounds, but on the notion that the universe must be mathematically harmonious and simple—a neo-Pythagorean view, derived possibly from the Platonic revival in fifteenth-century Italy, and foreign to inductive logic. The theory owed its acceptance largely to non-scientific reasoning, and to represent it as one of the major early successes of inductive thought is quite misleading.

The inductive method, then, did not come full-grown into existence ; it had to be puzzled out. A history of this process, and the confusions it engendered, might be very illuminating for the history of western thought.

7

To sum up : physics, if the foregoing account is reliable, has no philosophical implications.[1] If we consider it purely as a systematisation of observations by means of laws and theories, it has no metaphysical presuppositions either. If we consider it as a description of nature from a certain angle, it requires the metaphysical presupposition of order in nature. These conclusions have been reached from a consideration of the logic implicit in scientific method, which reveals that the formulation of empirical laws and of theoretical schemes would be logically invalid unless some such principle could be assumed. The critical study of scientific method is the key to the whole problem.

E. F. Caldin

---

[1] To affirm this is not to deny that physics may influence philosophy ; for it may call the attention of philosophers to aspects of the world that have become neglected.

THE PROBLEM OF THE TEMPORAL RELATION OF CAUSE AND EFFECT

1 *Introduction*

THREE views on the temporal relation of cause and effect appear possible, though four have been suggested. The fourth view, that cause and effect are ' contiguous in time ' seems to reduce to the limiting case of antecedence, but for the sake of completeness I have included the fourth supposition. The four proposed views are, therefore :

I Cause and effect are contiguous in time.
II Cause is antecedent to effect.
III Causes are at least partly contemporaneous with their effects : that is, a cause begins before its effect, but continues during at least part of the time occupied by the effect.
IV Cause and effect are coterminous in time.

I Russell says, in *Mysticism and Logic* :[1]

> Philosophers, no doubt think of cause and effect as contiguous in time, but this, for reasons already given, is impossible. Hence, since there are no infinitesimal time intervals, there must be some finite lapse of time τ between cause and effect. This, however, at once raises insuperable difficulties. However short we make the interval τ, something may happen during this interval which prevents the expected result.

In fact, however, philosophers are by no means all satisfied with the notion of contiguity in time, which seems to derive from Hume, but is not universally received in ancient or modern philosophy.

We find the following in Aristotle, whose argument is singularly like Russell's own :

> We cannot argue that because an event A has occurred, therefore an event B has occurred subsequently to A but still in the past—and the same holds good if the occurrence is future : cannot reason because, be the time interval definite or indefinite, it will never be possible to infer that because it is true to say that A occurred, therefore it is true to say that B, the subsequent event, occurred ; for in the

---

[1] Bertrand Russell, *Mysticism and Logic*, 8th ed., London, 1949

> interval between the events, though A has already occurred, the latter statement will be false.[1]

But this simply shows that Aristotle is in agreement on the inconvenience of supposing a time-interval, long or short. We require to know his opinion on the subject of contiguity in time, and we find that he considers the relation an impossible one :[2]

> It is evident, we may suggest, that a past event and a present process cannot be 'contiguous,' for not even two past events can be 'contiguous.' For past events are limits and atomic ; so just as points are not 'contiguous' neither are past events, since both are indivisible. For the same reason, a past event and a present process cannot be 'contiguous,' for the process is divisible, and the event indivisible. Thus the relation of present process to past event is analogous to that of line and point, since a process contains an infinity of past events.[3]

Kant is not so unambiguous as Aristotle, and it is, indeed, difficult to free him from the accusation of having accepted two contradictory hypotheses together ; for he accepts the notion of antecedent causation (of the antecedence of the cause), but when making his own analysis seems inclined to a theory of contemporaneous causation. It is at least clear, however, that he would reject the absurd formulation which Russell quotes from Baldwin. The following passage contains a clear rejection of Baldwin's 'whenever the first ceases to exist the second comes into existence immediately after' :

> Der grösste Theil der wirkenden Ursachen in der Natur ist mit ihren Wirkungen zugleich, und die Zeitfolge der letzteren wird nur dadurch veranlasst, dass die Ursache ihre ganze Wirkung nicht in einem Augenblicke verrichten kann. Aber in dem Augenblicke, da sie zuerst entsteht, ist sie mit der Causalität ihrer Ursache jederzeit zugleich, weil, wenn jene einen Augenblick vorher aufgehört hätte, zu sein, diese gar nicht entstanden wäre.[4]

Whatever effect such a statement may have upon the possibility of holding to antecedent causes in general, it certainly seems to dispose of the notion of 'contiguity.'

---

[1] *Post. Anal.* 95a, 34. Quotations in English are from G. R. G. Mure's translation in *The Works of Aristotle translated into English*, ed. W. D. Ross, Oxford, 1926 ; in Greek from *Aristotle's Prior and Posterior Analytics*, revised text and commentary by W. D. Ross, Oxford, 1949.

[2] '*ἢ δῆλον ὅτι οὐκ ἔστιν ἐχόμενον γεγονότος γινόμενον ; οὐδὲ γὰρ γενόμενον γενομένου· πέρατα γὰρ καὶ ἄτομα· ὥσπερ οὐδὲ στιγμαί εἰσιν ἀλλήλων ἐχόμεναι, οὐδὲ γενόμενα· ἄμφω γὰρ ἀδιαίρετα.*' *Post. Anal.* 95b, 3

[3] *Post. Anal.* 95–96

[4] *Kritik der reinen Vernunft* (1787). B. 248

We may, I think, with these authorities, reject the notion of 'contiguity in time' as an attempt to slip in an illegitimate third term between being compresent and being not-compresent in time. If it has any meaning, it must be either 'being compresent for a minimal duration,' or 'being as near in time as is possible without being compresent'; so that it disappears as a separate theory of the temporal relations of cause and effect.

II For the notion of the antecedence of cause to effect, Hume would appear to be the principal authority, though it is also accepted by Mill and, with other qualifications which almost or quite destroy it, by Kant. It is discussed, and not formally rejected, by Aristotle; but what value, if any, he attached to it, it is very difficult to say.

Hume's definition[1] is well known: 'We may define a Cause to be "An object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedency and contiguity to those objects that resemble the latter."[2]'

Mill accepts the notion of antecedence, but a very careful reading of his exposition will, I think, suggest that it was not necessary to his theory of causation. He was more concerned with the question of which of the necessary conditions of an event we should consider as the cause, than with that of the temporal relation of cause and effect. He wished to extend the notion of cause to *all* the necessary conditions, as his most personal definition of cause shows: 'The cause, then, philosophically speaking, is the sum total of the conditions, positive and negative taken together; the whole of the contingencies of every description, which being realized, the consequent invariably follows.'[3] This extension of the term is obviously legitimate, but I think it was unnecessary, and was only motivated by the excessive importance which Mill attached to causes. I shall discuss the question at length below, and so will say no more about it here.

It seems to me that the word 'follows' in the definition need not necessarily imply sequence in time. If all the necessary and sufficient conditions are present, the event should be taking place, otherwise I do not see how we can say that the conditions *are* sufficient.

---

[1] He gives, of course, a second definition, but it is not necessary to consider it here.

[2] *Treatise*, Book I, pt. III, sect. xiv

[3] J. S. Mill, *A System of Logic ratiocinative and inductive*, 9th ed., London, 1875, vol. I, p. 383

That Mill was not particularly interested in the temporal question is shown by the following :

> Whether the cause and its effect be necessarily successive or not, the beginning of a phenomenon is what implies a cause, and causation is the law of the succession of phenomena.[1]
> Whether the effect coincides in point of time with, or immediately follows, the hindermost of its conditions, is immaterial. At all events it does not precede it. . . . [2]

In spite of these admissions, however, Mill does generally speak of causation in terms of invariable succession of cause and effect :

> The Law of Causation, the recognition of which is the main pillar of inductive science, is but the familiar truth that invariability of succession is found by observation to obtain between every fact in nature and some other fact which preceded it. . .[3]
> To certain facts, certain facts do, and, as we believe, will continue to succeed. The invariable antecedent is termed the cause ; the invariable consequent, the effect.[4]

It is equally indisputable that Kant uses the language of antecedent causation :

> If I omit from the concept of a cause the time in which something follows upon something else in conformity with a rule, I should find in the pure category nothing further than that there is something from which we can conclude to the existence of something else. In that case not only would we be unable to distinguish cause and effect from one another, but since the power to draw such inferences requires conditions of which I know nothing, the concept would yield no indication how it applies to any object.[5]
> That everything which happens has a cause, is a law of nature. Since the causality of this cause, that is, the action of the cause, is antecedent in time to the effect which has ensued upon it, it cannot itself have always existed. . . . (' Das Naturgesetz : dass alles, was geschieht, eine Ursache habe, dass die Causalität dieser Ursache, d.i. die Handlung, da sie in die Zeit vorhergeht und in Betracht einer Wirkung, die da entstanden, selbst nicht immer gewesen sein kann.') [6]

---

[1] *Op. cit.*, vol. I, p. 397 [2] *Ibid.* [3] *Ibid.*, p. 377

[4] *Ibid.*, see also *op. cit.*, vol. I, pp. 387, 390, 392

[5] Kant, *op. cit.*, B. 301. Quotations in English, unless noted as my own translation, are taken from N. K. Smith, *Immanuel Kant's Critique of Pure Reason*, 2nd imp., London, 1933.

[6] Kant, *op. cit.*, B. 570

III and IV I shall not deal separately with the possibility of a partial overlap in time of cause and effect, for though it might be a way of reconciling the two statements that causes are before their effects, and that they are cotemporal with them, it does not seem to have been advocated explicitly. I shall also return to this particular question later on, so far as it needs to be dealt with. The important thing is to show that our authors have by no means excluded the possibility of contemporaneous causation.

Whereas Mill asks ‘ Does a cause always stand with its effect in relation of antecedent and consequent ? ’ [1], Aristotle says, ‘ But experience seems to show us that there are also causes distinct in time from their effects. Is this really so ? ’ [2]

Aristotle, indeed, is so much concerned in the *Posterior Analytics* with contemporaneous causation, that quotation is superfluous. It is sufficient to recall that the example of cause and effect which he renders classical by devoting almost half the treatise to its analysis (the eclipse of the moon), is, of course, an example of contemporaneous causation.

He not only objects to anterior causation on the grounds that any interval of time would make inference from cause to effect unreliable, but doubts if we could even argue from effect to cause. For, he says if we say that C (which now exists) implies an antecedent B, and B, implies a further antecedent A ; and therefore C implies A : shall we not, if we so argue, be obliged to find a middle term between C and B, and so *in infinitum ?*[3] Nevertheless, he does not say explicitly that anterior causation is impossible. His authority is, however, powerfully upon the side of contemporaneous causes as most important and most interesting for scientific purposes.

Hume argues against the possibility of contemporary causes, making the interesting claim, that if any are contemporary all must be, or rather, perhaps, that there can be no question of *partial* overlap, but that the cause must either be entirely before, or entirely with, the effect :

> 'Tis an established maxim, both in natural and moral philosophy, that an object, which exists for any time in its full perfection without

---

[1] Mill, *op. cit.*, vol. I, p. 395

[2] ‘ *ἐπὶ δὲ τῶν μὴ ἅμα ἆῤ ἔστιν ἐν τῷ συνεχεῖ χρόνῳ, ὥσπερ δοκεῖ ἡμῖν, ἄλλα ἄλλων αἴτια εἶναι, . . . καὶ τοῦ γίνεσθαι δέ, εἴ τι ἔμπροσθεν ἐγένετο ;* *Post. Anal.* 95a, 24.

[3] *Post. Anal.*, 95b. ‘ *ἢ ἀεὶ παρεμπεσεῖται διὰ τὸ ἄπειρον ;*

> producing another, is not its sole cause ; but is assisted by some other principle, which pushes it from its state of inactivity, and makes it exert that energy of which it is secretly possesst. Now if any cause may be perfectly co-temporary with its effect, 'tis certain, according to this maxim, that they must all of them be so : since any one of them, which retards its operation for a single moment, exerts not itself at that very individual time, in which it might have operated ; and therefore is no proper cause. The consequence of this would be no less than the destruction of that succession of causes, which we observe in the world ; and indeed, the utter annihilation of time. For if one cause were co-temporary with its effect, and this effect with *its* effect, and so on, 'tis plain there would be no such thing as succession, and all objects would be co-existent.[1]

Hume, then, is entirely averse from contemporaneous causation, but neither Kant nor Mill takes up this extreme position.

Kant goes so far as to say, ' Alle Veränderung ist also nur durch eine continuirliche Handlung der Causalität möglich, welche, sofern sie gleichförmig ist, ein Moment heisst.' [2]

> Now every alteration has a cause which evinces its causality in the whole time in which the alteration takes place. This cause, therefore, does not engender the alteration suddenly, that is, at once or in one instant, but in a time ; so that, as the time increases from the initial instant (a) to its completion in (b), the magnitude of the reality ($b-a$) is in like manner generated through all smaller degrees which are contained between the first and the last. All alteration is thus only possible through a continuous action of the causality which, so far as it is uniform, is entitled a *Moment.* The alteration does not consist of these *Moments*, but is generated by them as their effect.[3]

I have already quoted the beginning of a passage in which Kant remarks that most efficient causes in nature are contemporaneous with their effects. I must now, at the risk of becoming somewhat tedious, transcribe the rest of the passage, because it raises again the objection that a purely co-temporal notion of causation would make it impossible to distinguish which was cause and which effect. This is a difficulty which evidently troubled Kant a good deal, and as I think it is an unnecessary one, and shall endeavour to show this later on, I now quote the whole of the passage, though I shall not criticise it at this point.

> Here we must note carefully that it is a question of the order of time, not of the passage (*Ablauf*) of time. The time between the causality

---

[1] Hume, *Treatise*, Book I, Pt. III, Sect. ii

[2] and [3] Kant, *op. cit.*, B. 254

> (*Causalität*) of the causal agent (*Ursache*) and its immediate effect may be vanishing in quantity (and so the two may be together in time : viz, the causality and the effect), but the relation of the one to the other remains specifiable according to time (*der Zeit nach bestimmbar*). If I lay a ball of lead on a cushion and consider it as a cause of the indentation of the cushion, it is certainly contemporaneous with the effect. But I distinguish one from the other through the temporal relation of the dynamic connection of one with the other. For if I lay the ball on the cushion, there follows upon the smooth form of the cushion the indentation ; but if the cushion has (I know not why) an indentation, there does not follow upon that a ball of lead.
>
> And so the sequence in time is the only empirical criterion of the effect, with reference to the causality of the causal agent, which is anterior (*die vorhergeht*). A glass tube is the cause of the elevation of a column of water above the general level of the water in a larger vessel into which the tube is dipped ; is the cause, although both phenomena are together in time. For as soon as I draw up the column from the water in the larger vessel, something follows (*erfolgt*), namely the alteration of the water-surface from the horizontal form which it had in the larger vessel to the concave form which it has in the glass tube.[1]

It will not be necessary to add much to what has already been said with regard to Mill's position. When he is not explicitly concerned with the question of temporal sequence he sometimes makes a revealing slip. Discussing gravitation as a ' convenient fiction ' of men of science, he says that they can represent gravitation to themselves as ' exhausted by each effort, and therefore constituting at each successive instant a fresh fact, simultaneous with, or only immediately preceding, the effect.'[2] Now it is notorious that if one thinks of gravitation at all one thinks of it as acting throughout the time of its effects. The parenthesis is therefore put in to save precedence, which really has nothing to do with this example.

But perhaps the most interesting passage in Mill's work is that in which he seeks to explain why one necessary condition should have been singled out as *the* cause of an event, whereas the others were neglected. It seems to me that his argument proves the reverse of what is intended ; it seems to give sufficient ground *for* distinguishing one of the conditions from all the others. It also might be used to explain why causes seem to precede their effects ; because we do not lay *enough* emphasis on the difference between necessary 'causes' and necessary-and-sufficient causes :

---

[1] Kant, *op. cit.*, B. 248, 9 (my own free translation)

[2] Mill, *ibid.*, vol. 1, p. 382

> What, in the case we have supposed, disguises the incorrectness of the expression, is this : that the various conditions . . . were not *events* . . . but *states*, possessing more or less of permanency ; and might therefore have preceded the effect by an indefinite length of duration, for want of the event which was requisite to complete the required concurrence of conditions : while as soon as that event . . . occurs, no other cause is waited for, but the effect begins immediately to take place : and hence the appearance is presented of a more immediate and close connexion between the effect and the one antecedent,[1] than between the effect and the remaining conditions. But though we may think proper to give the name of cause to that one condition, the fulfilment of which completes the tale, and brings about the effect without further delay ; this condition has really no closer relation to the effect than any of the other conditions has.[2]

With the main thesis of this passage one may agree wholeheartedly, but it would perhaps have been better to say that what we *properly* call the cause, which is contemporaneous with and determines the time of the effect, is often a comparatively trivial condition which determines very little *but* the time of the effect, and is therefore frequently of very minor interest to science.

It really seems as though the theory of antecedent causation was a foreign body in the philosophies of Kant and Mill, an irritant around which they secreted nacreous material, but which they were unable to extrude.

In the attempt to refine the word 'cause' for use as part of a scientific vocabulary there has, perhaps, been some misunderstanding due to the implicit assumption, which no one would make explicitly, that a word can have only one precise legitimate meaning. Russell speaks of words as used in common speech as being like lumps of putty, with which it is easier to hit a target than with a smaller object like a bullet, but which allow less selectivity in the part of the target aimed at. This is true of many words, but there are clearly some which have more than one legitimate meaning according to their context, and for which no one precise meaning can be found, though it would be possible to refine the separate meanings separately.

The word 'cause' seems to show both the above properties at once in different 'dimensions.' If we take the word as referring

[1] But is it ?

[2] Mill, *op. cit.*, vol. I, p. 379

loosely to any necessary condition of an event, it is, I think, possible to show that it is most appropriate to one particular kind of condition ; namely that of which, Mill says, ' the fulfilment completes the tale, and brings about the effect without further delay.'

But it is clear also that the word ' cause ' is used with respect to temporal conditions in two distinct ways, neither of which can be declared more legitimate than the other. These two usages correspond to the notions of antecedent and contemporaneous causation with which we have seen the philosophers to be in labour.

Thus, if I die and leave money to a friend, I am undoubtedly the cause of his pleasure in spending the money, though I have ceased to exist by the time he is spending it. On the other hand it is equally legitimate to say that the cause of his pleasure is the actual spending of the money and the consumption or possession of the resulting goods, which are, of course, contemporaneous with the pleasure which is their effect.

Similarly—to avoid the suggestions that we are dealing with a difference between moral and physical causation—it is clearly proper to say that the cause of my feeling warmth on my face at sunrise is the sun ; though it is equally correct to say that I feel the sun's rays. The sun, as cause, is greatly anterior in time to the effect. The ' rays ' which I feel are contemporaneous with the feeling.

It is clear also that the remote cause has immediate effects of its own : I make my will when I make it, and the sun emits its ' rays ' when it emits them.

So far we are on familiar ground, but we already differ from Hume, who considers the cause to be ' contiguous ' with its immediate effect, and explicitly excludes the possibility of its being contemporaneous. It seems, however, quite indisputable that we have here two *radically* different usages : in the one, the cause is really anterior in time to the effect ; in the other it is at least largely contemporaneous.

It appears to me that the only matter of dispute which can arise in the above two examples is whether the cause is coterminous in time with its immediate effect, or only overlaps the effect considerably.

Before carrying this question any further, it will be well to deal with the question of ' invariable antecedence ', the notion of which is contained in Hume's definition. This notion clearly contains no formal contradiction, but it does not seem to harmonise with our experience. Whenever we call some antecedent a cause, we find that antecedent not to be invariable ; and whenever we call something invariable a cause, we find that it is not antecedent.

Let us take one of the few examples which Hume vouchsafes to us ; the movement of a billiard-ball. Let us also confine our attention for the moment to the initiation of its motion. What is the cause of this movement ? If we answer, the movement of another ball ; or the cue ; or that it was set in motion by one's hand ; it is obvious that these agents, considered as antecedents, are various. If, on the other hand, it is answered that something more refined is intended, that the invariable to which reference is made is the act of pushing or propelling the ball whose movement is considered as effect, then it is clear that this invariable is not antecedent, and by definition cannot be antecedent to the inception of the motion of the affected ball. There is no pushing unless something is pushed, and a thing is pushed just so long as something else is pushing it. That the effect really or apparently persists after the cause has ceased, introduces a problem which I hope to deal with later on ; but there seems to be no question but that the cause begins, and must begin, with the effect, and that no part of it can be antecedent to the effect.

Or let us ask, what is the cause of death ? Even to ask the question appears ridiculous, for there is an indefinite number of kinds of causes of death. Yet, if a cause be an invariable antecedent, there should be an event A which invariably precedes death, B. Or, if what is meant is not ' no A, no B,' but ' if A, then B,' there might be several causes of death, but they should, to satisfy the criterion, always be followed by death. Now this appears to be the case, but, as Aristotle and Russell have pointed out, as long as there remains an interval of time, there remains, at least theoretically, the possibility of the intervention of some prohibiting factor. But this, of itself, is not an insuperable difficulty, for nothing obliges us to suppose that the inference from cause to effect is a logical inference which must hold under all circumstances : it may be, as Hume and Russell assert, that we are concerned only with a general statement of what is generally the case.

But if we collect cases in which there seems to be an invariable sequence of cause (of death) and effect, we find such things as drinking cyanide, being shot through the heart, being decapitated, or falling into a furnace. But it is not the drinking, or the falling, or the shooting which is the immediate cause of death, but the combination of the poison with the substances of our body, the coagulation of our proteins by the heat, or the stoppage of the heart's action, all of which are (as is also the decapitation) contemporaneous with what we call 'death.'

Death, certainly, is a difficult subject to discuss, for it is not entirely clear what we mean by the word ; but if we take a simpler example, such as the boiling of a kettle, we can see how the invariable cause only *appears* to be anterior in time to the effect. In this case we happen to be interested in what is a relatively arbitrary stage in the continuous process of heating, in which it is manifest that cause and effect must be contemporaneous. Or, if we insist that boiling is not an arbitrary part of a process, we should have to find its invariable cause, and should have to mention, let us say, the escape of dissolved gases from a liquid. This would not only be contemporaneous with the effect, but would bring us to the position which Aristotle reaches in his analysis : the cause enters into the definition of the effect.[1]

It appears, then, that we may apply the word 'cause' to an antecedent ; but that if we require a 'cause' to be invariable, we shall find ourselves impelled towards something which is contemporaneous (though not necessarily coterminous) with the effect.

2 *Inductive Study of the Notion of Cause in Everyday Life*

The notion of causation seems to occur naturally in two sets of circumstances in everyday life :

I in productive arts, in which we are more or less aware of causes, and apply them ;

II in cases where some sudden change, or some unfamiliar conformation in the environment, or in ourselves, arouses our interest, and we seek a cause.

In production we do, of course, make our preparations and assemble our materials beforehand. It is a nice question to what extent the ordinary man would consider these as causal. Perhaps we must admit that we have here one of the legitimate extensions of the idea. Certainly medieval thinkers were familiar with the notion of 'material causation.'[2]

I think it safe to assume that the actual process of production would be considered as causal. Now in most, if not all, productive processes cause and effect are simultaneous : in all painting, chasing, planing and polishing ; in moulding, shaping, building and construction generally ; in the making of thread, weaving, cutting-out and making up of clothing ; in the cleansing and cooking of food : in mining and quarrying ; in ploughing, sowing, reaping, binding and stacking.

---

[1] Aristotle, *Post. Anal.*, inter alia : '*ἐν ἅπασι γὰρ τούτοις φανερόν ἐστιν ὅτι τὸ αὐτό ἐστι τὸ τί ἐστι καὶ διὰ τί ἔστιν.*' 90 a, 14.

[2] The notion is, of course, Aristotelian.

It is perhaps in the parts of farming in which human agency has least to do that we seem to find causes anterior to their effects. Thus the harvesting of corn is a remote effect of sowing. But here we have complex chains of causation, the consideration of which must be deferred to our next section where we hope to show what is the real nature of causal chains ; or, rather, to show how they are most naturally interpreted.

When some change in the environment, or some sudden change in ourselves is experienced, we often formulate a question in which the notion, and even the word, may be present : What's the cause of that ? Some sudden noise, some shadow, some local movement of the leaves of a tree, the grass of a field, or the surface of a pond ; or some unexpected feeling of discomfort or pain in ourselves ; all these may produce an attitude of attention and questioning which, if expressed in words, might receive this formulation. We should, I think, be surprised to learn that we *necessarily* suppose the cause we seek, in these cases, to be something anterior in time to the effect we have observed, or that the effect implies one kind, and only one kind of cause.

To these two classes of cases in which the idea of causation naturally occurs, we may find analogies in scientific investigation. On the productive side, we wish to ' try the effect ' of heat upon some substance, to discover the effect of bringing two substances together, to discover the effect of a drug, or to produce some known effect by the injection of a substance into an animal. In all such cases we seem to be dealing with contemporaneous causation.[1]

On the other hand, we are frequently confronted with ' effects ' of which we wish to discover the causes. A good example of this situation is provided by some considerable irregularity in an otherwise smooth curve. Here, again, we do not *necessarily* suppose that the cause we are seeking will be something which occurred before the effect. Thus if we find a sudden and remarkable change in the acceleration of growth, such as occurs in man at adolescence, we are very apt to suppose that it will be due to the activity of hormones which we shall suppose to be active during most or all of the time during which the general curve is departed from.

---

[1] The *injection* of a drug is, of course, anterior to the drug's effect ; but the injection is not the cause we are interested in. We suppose the drug to exert its effect by entering into combination with some substance in the body, or by acting as a catalyst. Both processes are, of course, unitary processes in time.

It is important to observe that what we are inclined to call ' an effect ' for which we seek ' a cause,' is not necessarily some change which we actually observe ' im Werden ' ; it may be some irregularity of an otherwise regular structure, of which a classical example would be a footprint in otherwise smooth sand.

3 *Analysis of Causal Sequence, and the Problem of Causal Chains*

We must now take our analysis of some simple cases of causal sequence a step further.

Suppose we strike a billiard-ball with a cue. Shall we call the cue the cause of the ball's motion ? Of course we can if we like, and since the cue existed before the event, the cause will be to that extent before the effect. But it is to be observed that the cue goes on existing after we strike the ball ; so that, by the same convention the cause is after the effect. What is of greater importance is that, though the statement ' this cue was the cause,' or ' this cue is the causal agent of the past event,' is a simple statement of fact, any statement about the cue before the event is extremely hypothetical. It seems then, that we are on much safer ground if we call the cue the cause *after* than *before* the event.

A much better convention would be to call the motion of the cue the cause, because, before the event, we cannot say that if the cue continues to exist it will strike this particular ball, or even with more than tolerable probability that it will strike any ball at all. But of the motion of the already directed cue we can say that *if* it continues unchanged (and the other circumstances of the experiment remain unchanged) it *will* be a cause of motion in this particular ball. But in neither case is there, strictly speaking, any causal event, or any cause of motion in the ball, until it is actually moved. This seems to me so obvious as to be hardly worth saying, yet it appears to be flatly in contradiction with the theory of universal antecedence in causes.

It seems quite clear that the cause begins with the effect, but it is by no means clear that the two terminate together. Yet this seems to me to be the case.

However smoothly continuous the motion of the ball may appear, we can distinguish in it two phases : a very brief one in which the ball is actually being struck, and a much longer one during which the motion is continued ' by inertia.' Moreover, when it is struck, it is, strictly, as much a cause with reference to the cue as the cue is in reference to it. The convention which considers the motion of the cue (or the cue itself) as the cause is justified because the change in the state of the ball due to the collision is much more striking than any visible effect either upon the cue or upon its motion.

But there is no disputing that we are here employing a convention.

It appears to me, then, that we should refine our terminology at this point by introducing some other term than ' effect' for the inertial motion of the ball. Fortunately no neologism is required, for we can legitimately call this part of the motion ' the consequence ' of the causal event.

We shall thus have a part of the motion of which the *duration* requires no cause. If now the ball strikes another, and that, after inertial motion, another, we shall have a causal chain of alternating causal and non-causal events. I do not know whether this is at all a novel suggestion, but it seems to me to be an account of what we actually suppose to be the case.[1]

It is clear that this would answer Hume's argument that simultaneous causation would make all events in the universe simultaneous and destroy time.

But another way in which we might account for chains of simultaneous causation is provided by the case of some object, let us say a book, which is pushed off a table and falls to the ground. Here the pushing and the being pushed are coterminous and, since the book is all the time in the earth's gravitational field, the last part of the push overlaps the first part of the free pull of gravitation, as the book is released from the support of the table.

An interesting parallel to this case is to be found in the example of the causation of night and day, which has presented a classical problem to those who define causes as ' invariable antecedents.' Just as, in the case of the book, we have a causal event taking place in the gravitational field of the earth, so, in that of night and day, we have the revolution of an opaque body, the earth, taking place in the ' light-field ' of the sun.

---

[1] I mean, of course, that this is what we think as ' common men ' living in a Newtonian Universe. Cf. R. G. Collingwood, *Proc. Aristotelian Soc.*, 24th Jan. 1938.

This reminds us of Mill's Permanent Causes,[1] and of Aristotle's Unmoved Movers.[2]

It is to be observed that we do not seek a cause of gravity (whether we are justified in seeking a cause of light I cannot say). With respect to the first law of motion, gravitation is a cause, because we postulate it to account for an ' irregularity ' : for the fact that the book does not continue its horizontal motion as the billiard-ball does ; and for the pathway of a projectile fired from a gun. Considered in itself, gravitation is a regularity, and this, it seems to me, is the reason we do not require it to have a cause ; because causes are invoked to account for irregularities with respect to some regularity. Thus, in so far as a stationary ball is in a regular state, we require a cause of motion in it ; but when it is in a state of regular motion, we require no cause for this. Similarly with static systems ; we are less inclined to ask for the cause of the form of a crystal than for the cause of some departure from regularity in it.

4 *Conclusions*

A causal event would appear to be a collision, combination, or resultant of processes.

By definition it must have at least one element which could be distinguished conventionally or, in some cases, arbitrarily, as *the* cause.

It does not follow that everything which could be called an event must have a cause.

All the elements of the combination may, as Mill pointed out, in one sense be considered as causes, in so far as any *conditio sine qua non* may be considered to be a cause.

On the other hand, it is easier than Mill supposed to assign the name of cause to one selected element, at last in the great majority of cases. In a case of exceptional simplicity, such as the collision of two particles of equal mass and equal velocity suffering, after the event, equal deflections from their original paths, it is impossible to assign a cause other than the collision itself, or to assign an effect other than the two deviations taken together. But in the majority of cases one process suffers less deviation than the other of the processes which by their collision constitute the event. Thus, in the case of a collision between a very small and a very large particle, the smaller suffers greater deviation of its path than the larger. Or one conformation (considered as a type of ' process ') suffers greater deformation than another, as in the case of Kant's cushion depressed by the ball of lead. There is probably no case in which the causal agent suffers no deflection or deformation (' patients are always agents ; in a great proportion, indeed of all natural phenomena, they are so to such a degree as to react forcibly on the causes which acted upon them ') [1] ; so that there will probably always be reciprocity of causation, and the process considered as cause can also be considered as receptive of an effect.

---

[1] J. S. Mill, *A System of Logic*, 9th ed., London, 1875, vol. 1, pp. 398–399

[2] I understand it is now thought that Aristotle supposed there to be more than one.

Another criterion of *the* cause is suggested by Mill himself : ' that one condition, the fulfilment of which completes the tale, and brings about the effect without further delay.'[2] Mill is, of course, fully justified in considering this a convention. If we suppose an explosive mixture of four components, any three of which are insufficient to produce the effect, it is clear that any one which happens to be added last will be *the* cause of the explosion. As long as the convention is understood, there is no reason why it should not be retained.

The two criteria are independent, and consequently might be in conflict in some particular case. If one strikes a glass vessel against a stone, both the motion of the vessel and its conformation suffer the greater change, though the movement of the vessel is the last condition added. We simply have to employ the convention which happens best to suit the particular case.

The question of what is the effect is to some extent complementary to that of what is the cause. It does, however, present some features of special interest. In the case of an explosion, the explosion considered as effect is the same thing as the causal event. And, as pointed out above in the case of a collision of identical particles, the collision itself is both cause and effect ; that is, if we wish to specify *one* cause and *one* effect. But in general we can find some process which can be considered as receptive of an effect. Kant has endeavoured to distinguish cause and effect in his example of the cushion by pointing out that if a leaden ball is laid on a cushion, a dint follows in the cushion, whereas if the cushion ' for some reason ' has a dint, no leaden ball follows. But whatever heavy body makes the dint persists after the causal event, and is just as natural a consequence of the causal event as is the depression in the cushion. Nevertheless, there *is* a sequence of events, but the sequence is not that of cause and effect (which are simultaneous as Kant observes), but of smooth-surface-of-cushion followed by indented-surface.

---

[1] Mill, *op. cit.*, vol. 1, p. 388

[2] *Ibid.*, p. 379

Possibly this was what was in Kant's mind when he showed how the concave water-surface in the glass follows upon the plane surface of the water in the larger vessel, but I do not think he makes the point very clear.

In considering effects, we have to give our attention to the problem of consequences : are we justified in making a distinction here? Common speech makes no such distinction, so that it cannot be disputed that there is a sense in which effects (including consequences in effects) persist after their causal events ; and to this extent the effect is after its cause. But even if we make this concession to common speech, the effect begins with the cause, and so it is not *totally* after its cause.

But it appears to me that if we are to refine our terms at all we must allow this distinction, for it is the lack of it which has been partly responsible for the confusion over the temporal relations of cause and effect. In the case of the cushion the distinction is particularly clear, for if the cushion be sufficiently elastic the effect will terminate with the cause : its persistence requires an explanation in the ' properties ' of the cushion. Mill clearly perceives the importance of consequences, though he does not distinguish them from effects :

> There is a case of causation which calls for separate notice . . . It often happens that the effect, or one of the effects, of a cause, is not to produce of itself a certain phenomenon, but to fit something else for producing it. . . . Physiological agencies often have for the chief part of their operation to *predispose* the constitution to some mode of action, . . . We must, therefore, . . . include among the effects of causes the capacities given to objects of being causes of other effects.[1]

The theory of antecedent causation has been a fertile source of paradoxes : there is the classical problem of day and night, and more recently Russell has proposed his two paradoxes, namely, ' mnemic causation,' and what one might call ' the hooter paradox.' Russell has also shown how troublesome the theory is if we actually attempt to apply it in science, and he remarks : ' As a matter of fact, such sequences as were sought by believers in the traditional form of causation have not so far been found in nature.'[2]

If, however, we think of a causal event as a unitary event, of which both ' cause ' and ' effect ' are conventionally distinguished elements, all the troublesome paradoxes disappear ; moreover, we have, as I have attempted to show, a notion of causation which is not at all difficult to reconcile with the way in which we actually use the words ' cause ' and ' effect ' in every-day life.

---

[1] Mill, *op. cit.*, vol. I, p. 389

[2] Bertrand Russell, *Analysis of Mind*, 5th ed., London, 1949, p. 94

It appears to me that the theory of contemporaneous causation also lends itself to an easy transition from every-day usage to a more scientific application.

Contemporaneous cause and effect lend themselves readily to expression in terms of a function, in which cause is taken as the independent variable, and effect as the dependent. For example, let us take the length of a rod of metal considered as a function of temperature (if temperature seems too abstract to rank as a cause, we can take it in this case as measuring energy ; though we could not do so in all cases, of course). If we suppose the heating, as cause, to precede the expansion, as the effect, we can, of course, express the relation as a function, abstracting from the temporal sequence. But since temporal sequence will not find expression in the function, the function will be a better expression of the fact if no time interval actually exists. Now this appears to be the case. It is as and when the rod ' receives a certain quantity of heat ' that it expands, not *after* the causal event of heating. Also, we commonly suppose the expansion to be the expression of atomic or molecular movements which are contemporaneous with the expansion.

If we think of cause and effect as elements or aspects of one process, it follows that an effect must have a cause which is coterminous in time with itself ; but this does not exclude the possibility of its having remote causes which precede it in time, nor does it exclude the possibility of causal chains, as I have argued above. The notion is therefore compatible with both the ways in which we do in fact think of causes in every-day life.

It appears to me that to think of a causal event as a combination of processes allows us to give causation its proper place in science. In so far as causal events are repeatable, they themselves exemplify ' laws.' But in so far as they are, with respect to their component processes, irregularities, they cannot be the ultimate objects of science, which aims at the discovery of regularities as such.

Causation, then, seems to occupy a position intermediate between those assigned to it by Mill and Russell respectively : it is hardly to be dismissed as useful only ' in the infancy of a science,'[1] nor can it be considered as 'the main pillar of inductive science.'[1] It is simply a way of considering phenomena which is not ultimate or absolute, but which does appear to yield valuable results at least in the intermediate stages of the development of a science.

---

[1] Bertrand Russell, *Mysticism and Logic*, 8th ed., London, 1949, p. 194

But with such general questions I do not pretend to deal. I have tried in this paper only to show that a theory of causes which considers them as contemporaneous with their immediate effects is perfectly compatible with common sense, removes some of the troublesome paradoxes to which the theory of antecedent causation gives rise, and lends itself more readily than the latter theory to refinement for the purposes of exact science.

J. S. WILKIE

---

[1] J. S. Mill, *A System of Logic*, 9th ed., London, 1875, vol. 1, p. 377

SOME ASPECTS OF THE SEARCH FOR INVARIANTS

1 *Introduction*

SINCE the terms 'Invariant' and 'Invariance' are used somewhat loosely in physics, we will start by giving a definition covering their use in the present paper. '*By invariance, we mean the constancy of some complex entity or object of thought while certain of its constituent elements or associated quantities undergo change.*'[1]

Historically, the search for invariants must have arisen as a result of the development of exact measurement. Such quantities as length, volume, weight, etc. of individual objects or of samples of materials, were important for assessing values. Dependence on such subjective concepts as 'full measure, pressed down and shaken together,' would no doubt involve much argument, leading finally to carefully specified definitions of quantity and conditions of its measurement.

Later, the importance of quality of materials as distinct from quantity of samples made itself felt and the concept of 'physical property' was gradually developed. *Physical properties are now generally defined as scalar quantities, invariant to changes of axis and, following the establishment of empirical laws, also to changes in as many other variables as possible, certainly in all those included in their definition.*

The simplest type of physical property owes its invariance only to the homogeneity of the material : examples are density and surface tension. As an example of more complex properties, we may take the elastic coefficients, whose invariance depends on a convenient interaction of repulsive and attractive inter-atomic forces.

Many cases are known in which physical properties are found experimentally not to be invariant to changes in their dimensional components, e.g. the anomalous properties of certain dielectrics, behaviour associated with the work-hardening of metals, the anomalous viscosity of complex liquids, failure of Hooke's law of elasticity for concrete, etc. Moreover, Grunberg and Nissan[2] doubt whether any so-called true fluids retain an invariant viscosity for high rates of shear, and Zener and Hollomon[1] claim that ' no real metal possesses an ideally elastic region '.

---

[1] H. V. Craig, *Vector and Tensor Analysis*, New York, 1943, p. 7

[2] L. Grunberg and A. H. Nissan, *Nature*, 1945, **156**, 241

Moreover, if so-called properties are not truly invariant to changes in their constituent elements, materials will be found to have ' properties ' which depend on the methods of measurement ; for example, Garner and Nissan[2] have recently shown that certain materials have invariant viscosities when that property is measured by timing the fall of a metal sphere through the liquid ; but that, when the liquid is flowing through a tube, the viscosity varies with the rate of flow.

Dingle[3] has rejected the whole concept of ' physical property ' except for convenience in other than fundamental considerations. What is fundamental is the result of some specific measurement. If different processes of measurement give the same or approximately the same ' properties ', so much the better for the physicist, but he must not count on such happenings nor be perplexed when they do not occur.

There are difficulties, however, in the use of even such ' operational ' definitions. For example, however carefully we try to perform the same operation on different materials whose properties we wish to compare, the reaction of the materials may be so different as virtually to change the nature of the operation itself (see Scott Blair and Baron[4]). But in spite of these difficulties, operational definitions have many advantages over the more rigid concepts of invariant properties.

In earlier times, physicists, seldom working in industry, could safely ignore complex materials, unusual conditions, or comparatively small anomalies as the case may be : today, some method of classifying them must be found. Not infrequently in such cases, the physical property is still quoted with some specification of arbitrarily selected defining conditions. It must be remembered, however, that, except as an analogy with the behaviour of simpler systems, a physical property has really no significance unless it is invariant to changes in its defining elements, since it is only through such an invariance that it comes to be isolated as an independent concept.

If there are no invariant ' physical properties ' by which to describe the kind of behaviour we wish to study, what other kinds of invariance or partial invariance may we hope to find or what other kind of concept can we usefully define ?

[1] C. Zener and J. H. Hollomon, *Journal of Applied Physics*, 1946, **17**, 69

[2] F. H. Garner and A. H. Nissan, *Nature*, 1947, **170**, 329

[3] H. Dingle, *Philosophical Magazine*, 1949, **40**, 94 (And other papers in the same series.)

[4] G. W. Scott Blair and M. Baron, *Nature*, 1949, **164**, 148

We will consider three ways of approaching this problem: (a) The Analytical Method; (b) Statistical treatments such as Factor Analysis applied to the results of empirical tests; (c) The Principle of Intermediacy and the Theory of Quasi-properties.

2 *The Analytical Approach*

In some cases there is evidence and in others it is merely assumed that the reason for the apparent lack of invariance of the usual properties in a complex material is that there is an interaction of a number of units each of which itself has simple physical properties. In rheology [1] the complex behaviour of such materials is often accounted for by means of models consisting of elastic, viscous, static-frictional and other elements connected in series and in parallel, and in electricity very similar models are used. In certain cases these postulated mechanisms can be at least partially linked with molecular processes: in others, no such molecular picture is attempted and the proposed model is purely an analogy.

When a single elastic element is attached to a single viscous element in series, we have the case long ago described by Maxwell in which the stresses in a strained system are gradually relaxed in an exponential manner: i.e. the natural logarithm of the stress is reduced at a constant rate, the reciprocal of which rate, having the dimensions of a time, is known as 'the relaxation time' and is given by the ratio of the viscosity of the viscous element to the modulus of the elastic one.

In many real materials, however, as Maxwell himself knew, this idea of a single relaxation time is a gross over-simplification and the concept of an infinite number of relaxing Maxwellian units now constitutes a well-known feature of the classification of complex materials. In electricity, a similar treatment has been used to explain the distribution of time constants.

If a dielectric be considered as built up of a heterogeneous structure such that the ratio of conductance to dielectric constant differs for the different components, some of the complex phenomena of 'absorption' in constant fields and of power losses in alternating fields can be explained. As Hartshorn [2] puts it: 'It is obvious that by taking a sufficient number of terms any experimentally found curve can be represented and thus by assuming sufficiently heterogeneous structure almost any experimental curve can be explained by this theory'. Although in rheology, as in electricity, certain relatively simple systems can be adequately explained in terms of a very few components, Hartshorn's remark would be equally applicable to the former science.

---

[1] The science of the deformation and flow of matter.

[2] L. Hartshorn, *Journal of the Institute of Electrical Engineers*, 1926, **69**, 1152 (For a more advanced and modern account of dielectric problems, see H. Fröhlich, *Theory of Dielectrics*, Oxford, 1949.)

Recent work in rheology has sometimes postulated quite small numbers of elastic and viscous elements, but both these groups of elements have themselves complex properties—the former do not have an invariant simple ratio of stress to strain nor the latter a constant ratio of stress to rate of flow (viscosity). In some cases, the complex type of behaviour of the elements is itself explained on the basis of some theory (e.g. the Theory of Rate Processes). This explanation in terms of models which are themselves complex, may be regarded as of intermediate value, as, in certain cases, are the non-orthogonal factors discussed in the next section.

3 *Empirical Tests and Factorial Analysis*

There are so many cases in industry where it is of urgent practical importance to describe the behaviour of materials which do not appear to have 'invariant properties', that complicated batteries of empirical tests have frequently been devised. Each test gives a reading of a pointer on a scale of some kind. All that is required of such tests is that they should be 'replicable' when carried out on replicate samples of material (they are seldom 'reproducible' on the same sample since the process of testing changes the sample) and that they should give results which do not depend appreciably on the skill or personality of the experimenter. It is by no means always worth while to ensure simple physical conditions and the pointer readings obtained cannot be converted into properties which are invariant to changes in conditions, or at least no function is known which would render them invariant.

In such cases, every single sample must be described by as many magnitudes (n) as there are pointer readings. A study of the correlation coefficients between large numbers of samples in the tests ($1 . 2 . . . n$) will, no doubt, eliminate some tests as unnecessary and may suggest where there are gaps in the test battery.

Any treatment involving the use of correlation coefficients loses its effectiveness when applied to non-linear relations; but experience has shown that the loss of information due to non-linearity is seldom serious.[1]

A not dissimilar situation has arisen in psychometry and data from large numbers of tests are being increasingly reduced to order and meaning by the use of the various systems of factorial analysis.

Factorial analysis has been developed as a kind of modern counterpart of the now discarded ' Faculty Psychology ', factors being much more fluid concepts than the older faculties and types. Its development has been gradual and varied and has grown out of the use of correlation coefficients as a means of comparing tests. F. Galton, K. Pearson and C. Spearman were largely responsible for the earlier stages of this development (see Burt [2]). They developed the idea that mental ability could be described in terms of two factors (or more correctly, two kinds of factors), one ' general ', which is required in some measure for all kinds of activities, the other, the specific factors required for each individual type of activity.

Later, the theory was modified to include ' group factors ', common to some but not all tests, and a fourth type to include both errors and incidental variations depending on the circumstances when the test is performed.

It is now generally agreed that any such classification is arbitrary. What appears as a general factor in the analysis of one set of experiments may include a group factor in a later set in which new tests have been included; and a ' specific ' factor ceases to be specific as soon as more tests can be found which include its measurement in some degree.

Although psychometrists have themselves suggested that such methods might profitably be applied to the classification of empirical physical tests (e.g. Thurstone [3]), it is only very recently that preliminary attempts have been made by Harper and Baron [4] for rheological tests on cheese, and by Harper, Kent and Scott Blair [5] for both rheological and electrical tests on certain plastics.

The ultimate purpose of factorial analysis is to reduce the number of magnitudes by which samples must be described from n (the number of tests) to m ($< n$) factors which are normally independent of one another. In practice we seldom describe individual samples in this way, being content to study only the relationships between tests.

---

[1] The effect of non-linear relations is to increase the error, not to change the nature of the factors.

[2] C. Burt, *British Journal of Educational Psychology*, 1949, **19**, 100

[3] L. L. Thurstone, *Multiple Factorial Analysis*, Chicago, 1945

[4] R. Harper and M. Baron, *Nature*, 1948, **162**, 821

[5] R. Harper, A. J. Kent and G. W. Scott Blair, ' The Application of Multiple Factor Analysis of Industrial Test Data ', *British Journal of Applied Physics*, 1950, **1**, 23

Diagrams or models can be made in which straight lines join an origin to points representing the individual tests. These lines (' test vectors ') are really projections on to an m-dimensional space from a space having a number of dimensions greater than m ; and the ' communalities ', which are defined as the squares of the test vector lengths, indicate the extent to which the analysis in the m dimensions of the model accounts for the variance of each test.

The correlation coefficient between any pair of tests is represented by the scalar product of the test vectors, i.e. the product of their lengths and the cosine of the angle between them.

In order that different tests may be compared with one another, a dimensionless scale must be found. This is best done by scaling unit distance in every case to represent the standard deviation of the test.

Some of the variance will be specific to individual tests and not common to others and this, together with errors, has to be discarded if we are to make $m < n$. This is effected by replacing the unities in the diagonal cells of the matrix of correlation coefficients by the communalities, to give what is called the ' reduced matrix '.

Drawings can be made for cases where $m = 2$ and solid models when $m = 3$; beyond this, the use of vector models is not applicable though perhaps some kind of tensor models, partly analogous to the *C.I.E.* charts showing just noticeable differences of chromaticity, might be invented. In all such cases, of course, the individual samples are not represented on the diagrams but only the tests, in terms of their standard deviations.

For those who are familiar with the terminology of matrix algebra, the reduction in the number of dimensions from n to m is expressed by the equation

$$R = F . F'$$

where R is the square ' reduced ' matrix of correlation coefficients between the tests, F is the common factor matrix and F′ its transpose. F and F′ may be represented by rectangles whose divergence from squareness represents the amount of saving in dimensions which the analysis has been able to effect.

Although such a method analyses a matrix of correlation co-efficients, this is really a short cut to an analysis of the matrix of measurements, which is the ultimate object of factorial analysis.

Factors themselves may thus sometimes be regarded as ' ideal tests ', or to be more exact, as perfectly correlated with ideal tests, since we do not know for certain that equivalence in discriminatory properties is necessarily the same as identity. Theoretically, there would seem to be no good reason why this method should not be transposed [1] so that a matrix of correlation coefficients between samples in the different tests would be factored. In psychometry, such transpositions have a real importance: in physics it is rare that enough tests are done on the samples for the method to be practicable.

In the most usual method of factor analysis, the first axis of the diagram or model is drawn through the centroid of the points (analogous to the centre of gravity) which represent the tests, and other axes are added at right angles. This method ensures that the first factor accounts for the greatest possible amount of the variance and, with each new factor, the maximum amount of residual variance is extracted; but there is no reason why other sets of orthogonal axes should not equally well describe the relations between the tests. The configuration of the test vectors is invariant to changes of axes and various criteria have been proposed for the selection of the best axes.

In certain cases these configurations are most simply described in terms of non-orthogonal axes. Non-orthogonal solutions are sometimes regarded as intermediate solutions, the non-orthogonal factors being themselves factored to give so-called ' second order factors ' (see above).

It must be stressed that the number, nature and order of importance of the factors depends to some extent on the population of tests and samples selected for analysis. To be useful, however, the factors, though not themselves invariant, should have a reasonable ' stability ', i.e. should appear in recognisable form in the analysis of various sets of experiments.

Readers who require more than this very brief summary of factorial analysis are referred to text-books by Burt [2], Thomson,[3] and Thurstone (loc. cit.).

---

[1] This does not imply that the new factor matrix is the transpose of the old; indeed only in special cases is there any clearly established connection between them.

[2] C. Burt, *The Factors of the Mind*, 1940

[3] G. Thomson, *The Factorial Analysis of Human Ability*, 1948

4 *The Principle of Intermediacy and the Theory of Quasi-properties*

We will now consider an altogether different approach to the problem: one which has been developed mainly by the writer and his colleagues.

A physical property has definite dimensions, which may be derived from its definition in any given dimensional system. For example, a viscosity is defined as a stress (force per unit area = $ML^{-1}T^{-2}$ in the usual MLT system) divided by a rate of shear (T^{-1}) and therefore has the dimensions $ML^{-1}T^{-1}$. If we prefer to define our dimensions in other terms, e.g. force, length, time, or force, mass, length, time, the dimensions of viscosity will, of course, be correspondingly different; but, whatever system we use, the dimensions will remain the same for all viscous materials within that system. The defining of physical properties is, indeed, mainly concerned with finding experimentally some appropriate powers of (normally) mass, length and time which will give an invariant grouping, thus establishing a physical law.

The same groupings are not invariant for all materials. For instance, in considering the displacement of bound electrons or dipoles in a condenser, the invariant is the dielectric constant: in the free flow of electrons along a wire, it is the conductance. Real materials are neither perfect dielectrics nor perfect conductors. In an elastic solid, the stress divided by the strain is invariant; in a fluid, the stress divided by rate of straining. A reasonable number of real materials approximate in behaviour to these prototypes but many complex materials, now too important to ignore, fail between them. In our rheological example, what is found to be invariant is the stress divided by some entity intermediate between the deformation and its rate of change.

The writer has elsewhere pointed out[1] that the concepts of whole-number differentials of length with respect to time, such as velocities and accelerations, are constructs which depend on our definition of the equality of time units. This equality is ultimately defined in terms of the superposably equal distances which light or alternatively material bodies uninfluenced by gravitational fields, traverse in the periods to be compared.

The whole of Newtonian physics is based on this conception and we are far from proposing to use a different time scale to study the

---

[1] G. W. Scott Blair, B. C. Veinoglou and J. E. Caffyn, *Proceedings of the Royal Society* A, 1947, **189**, 69

Q

behaviour of different materials; but, if we rightly insist on using Newtonian time to describe complex systems, we cannot expect always to find whole-number differentials in our invariant groupings. Such complex systems have their own simplest time scales which do not depend on bodies uninfluenced by gravitational fields but on the very complex interaction of their molecular components.

The fractional differential equations derived by the writer and his colleagues lead, on integration, in the simplest case to power laws and in more complex cases to series equations containing power terms. It is remarkable how many complex rheological systems obey simple power laws relating stress, strain and time[1] and how well even the most complex systems can be accounted for quantitatively by the series equations. In a few cases fractional differentiation with respect to stress as well as to time is needed [2] but the number of parameters required to describe the most complex systems is much less than when an analytical treatment by means of exponential terms is applied.

The use of simple power laws to describe anomalies in dielectrics and in rheology is long established (see Hartshorn, loc. cit., and Leaderman[3]), but the theoretical significance of such treatments has been largely overlooked.

There have been some criticisms on dimensional grounds, since, for example, the constant in our rheological power-equation analogous to viscosity and elastic modulus (known as the 'intensity factor') has dimensions which lie between those of these two properties, i.e. $ML^{-1}T^{-2+k}$ where the value of k differs for different materials or groups of materials.

Scott Blair and Veinoglou (loc. cit.) have shown that for quite a large group of plastics there is an invariant having the dimensions $ML^{-1}T^{-1 \cdot 8}$ which may be regarded as a true physical property for this group of materials alone. We cannot compare the behaviour of this group with that of some other plastic having a different value of k solely in terms of the intensity factors, but must always associate them with the magnitudes of k and of the (fractional) exponent of stress if this is other than unity. This is quite unobjectionable on dimensional grounds, as Dingle (loc. cit.) has recently shown. Although the dimensional symbol represents the process of measurement and is the same for all applications of the process, and also the individual symbol must allow that physical relations be independent of the units of measurement, the form of the symbol need not be the same to meet these two requirements. The whole question rests on what kind of entity is taken as the ' dimensional symbol '.

---

[1] G. W. Scott Blair and B. C. Veinoglou, *Journal of Scientific Instruments*, 1944, **21**, 149

[2] G. W. Scott Blair and J. E. Caffyn, *Philosophical Magazine*, 1949, **40**, 80

[3] H. Leaderman, *Creep of Filamentous Materials*, Washington (D.C.), 1943

We have been at some pains to find a suitable nomenclature for our proposed treatment. The principle which recognises that the behaviour of real materials often lies between that of the dimensionally simple prototypes has been called the ' principle of intermediacy.' The Gestalt-like group of magnitudes involving the dimensionally complex intensity factor and a number of quality factors which are numerics, forms a unity for classificational purposes which cannot be divided into ' components ' while retaining its significance. The obvious term ' group ' is unsuitable on mathematical grounds and a special use of the word ' multitude ' has been coined to describe such unities.

Until recently, this approach has been used only by the writer and his colleagues in the formulation of rheological behaviour but there are now indications that a very similar treatment is being implicitly applied in quite a different field.

The surface tension of liquids is not invariant to temperature changes and van der Waals long ago derived a law, from thermodynamic considerations, involving a small power which should be, but is not in fact, constant for all liquids. McLeod proposed an empirical relation, later found to be deducible on statistical grounds, by which the difference in density of vapour and liquid was claimed to be proportional to the $\frac{1}{4}$ power of the surface tension. This led to Sugden's parachor, an invariant by means of which molecular volumes of atoms, molecules, radicals and types of chemical bonds could be additively compared, not at a constant temperature, but at constant surface tension, corresponding to approximately equal degrees of self-compression by intermolecular cohesional forces.[1]

In a paper by Ferguson and Kennedy,[2] the significance of which has recently been re-emphasised by Telang,[3] it was pointed out that the power in the parachor, defined as $\frac{1}{4}$, is not in fact independent of the substances studied. The ' revised parachor ', which has an exponent which differs for different substances, requires for its specification two inseparable magnitudes: one, an intensity factor and the other a power which varies from about 0·2 to 0·3. The exponent in the original van der Waals equation, as well as those in a number of invariants derived from the parachor (e.g. the rheochor[1] and the thermochor) are likewise affected (see Scott Blair[2]).

---

[1] N. K. Adam, *The Physics and Chemistry of Surfaces*, Oxford, 2nd ed., 1938

[2] A. Ferguson and S. J. Kennedy, *Transactions of the Faraday Society*, 1936, **32**, 1474

[3] M. S. Telang, *Journal of the American Chemical Society*, 1949, **71**, 1883, 1898.

It is clear that this new treatment does not lead us to any fundamental knowledge about the molecular structure of materials and this fact has sometimes formed the basis of criticisms; but we maintain that, despite the great contributions which molecular and atomic theories have made to physics, the phenomenological behaviour of materials forms a perfectly legitimate part of the work of the physicist whether it contributes or does not contribute to these particular theories.

The use of the term 'intermediate' has given rise to some misunderstanding. If we say that the expression $d^\nu l/dt^\nu$ is *intermediate* between dl/dt and $l = d^0 l/dt^0$ so long as $1 > \nu > 0$, we do not mean that the *values* are necessarily intermediate but that the *operation* is intermediate (for further discussion on these points see Scott Blair[3] and Scott Blair and Caffyn, loc. cit.).

Perhaps the most hopeful indications of the generalisation of these concepts to other branches of knowledge lie in the direction of psychophysics. Harper[4] has recently shown that a value of ν can be found which accounts quantitatively for some of the curious anomalies observed by Scott Blair and Coppen[5] in studying the subjective assessing of firmness of qualitatively different materials by handling, though it has not yet proved possible to determine the value of ν by independent experimental means.

It is clear that the 'properties' described by multitudes do not qualify to be called 'physical properties' according to our original definition.

Some physicists have no objection to an extension of the definition but, since others do object, we have preferred to coin the term 'quasi-property' to describe such entities.

Quasi-properties describe processes and, for many materials, any process that will give pointer readings will appreciably change the nature of the material while the readings are being taken. Quasi-properties thus differ in their degree of invariance to the process of measurement. In some cases, the values of the 'components' of the multitudes are much the same when widely different processes are used for the test: in others, they differ considerably. Quasi-properties at the top of this hierarchy, which resemble most closely classical 'physical properties', are much more useful for classifying materials than those which are strongly dependent on the method of measurement. The existence of the latter type may be deplored but it can neither be denied nor profitably overlooked.

---

[1] J. Newton Friend, *Nature*, 1942, **150**, 432

[2] G. W. Scott Blair, *Journal of the American Chemical Society*, 1949, **71**, 2950

[3] G. W. Scott Blair, *Measurements of Mind and Matter*, London

[4] R. Harper, *American Journal of Psychology*, 1947, **60**, 554

[5] G. W. Scott Blair and F. M. V. Coppen, *American Journal of Psychology*, 1942, **55**, 215; *ibid.*, 1943, **56**, 234

The multitude is a new kind of non-tensorial unit and there is, of course, a danger that if its nature is not understood, its misuse may lead to logical inconsistencies. This is the only sense in which we believe that the theory of quasi-properties is subject to criticism on dimensional grounds.

5 *A Comparison between the Principles Involved in Factorial Analysis and in the Theory of Quasi-properties*

These two methods of approach have much in common.

(*a*) They are both attempts to find entities which shall be as nearly invariant as possible, for systems which do not show invariant physical properties of the classical kind. In the latter treatment, invariance is found provided that the selected powers of the variables concerned in the definition are specified. In the former, configurations of test correlation co-efficients are found to have stability and durability, and, qualitatively, the same factors should be forthcoming from them, i.e. there should be a considerable measure of invariance to changes in tests and materials. Attempts may well be made to select tests with a view to maximising this invariance.

(*b*) The 'multitude' is in the nature of a Gestalt. It consists of more than the sum of its parts (see Scott Blair [1]). Materials cannot be compared in terms of any 'part' of the multitude since the parts are not components in the usual sense of the word. The vector configuration is likewise a 'whole'. If one test-vector is withdrawn the whole configuration readjusts itself ('field-property'). Pairs of vectors which form 'doublets' cannot be divided into two independent members having separately definable loadings.[2]

---

[1] G. W. Scott Blair, *Journal of Colloid Science*, 1947, **2**, 21

[2] The 'loading' may be described as the correlation coefficient between the test and the factor.

(*c*) In the factor models, if the points representing the original samples are included and not only those representing tests (i.e. ideal samples which give readings of one standard deviation greater than average), such points may be defined in terms of polar co-ordinates by their distance from the origin, which is a kind of intensity factor, and the angles which they make with the axes, which describe their degree of intermediacy between those orthogonal factors which may be regarded as ' ideal tests '. But the distance from the origin is completely defined by the x, y, and z Cartesian co-ordinates whereas it is not the *value* of the intensity factor of the multitude which is defined by the magnitude of the exponents, but only its *dimensions*. This follows from the fact that the tests, however disparate dimensionally, are all plotted in the factor models in terms of ratios of standard deviations. The intensity factors in the factorial models are dimensionless, thus having a considerable advantage over their counterparts in the theory of quasi-properties.

(*d*) Models have been made to illustrate quasi-properties in which the intensity factors of each material are plotted vertically against the two dimensionless exponents. In these models, ideal behaviour is represented by the corner of a square.[1] Such models bear some resemblance to factor models in which the ideal tests form the rectangular axes of the model. The factor model has the advantage of being in metric space whereas the quasi-property model is in non-metric space.

(*e*) The use of non-orthogonal factors has not been paralled in the theory of quasi-properties. It bears some resemblance, however, to the use of non-ideal elements in the models proposed in the rheological applications of the Theory of Rate Processes (see Halsey[2]). In both cases such models may often be regarded as marking a temporary halting place rather than as a permanent solution. The anomalous elements must be explained in terms of ultimately simple behaviour and the non-orthogonal factors may themselves be factored.

It is tempting to seek for more direct connections between properties, quasi-properties and factors but it must be confessed that, in so far as the last two are concerned, it is hardly likely that any simple connection will be found.

In the case of systems simple enough in structure to show invariant properties of the kind used in classical physics, no doubt a factorial analysis of data from a set of empirical measurements would produce factors corresponding to the physical properties. In the factorial analysis of data obtained from complex systems, however, we start with a set of pointer readings which are not invariant with respect to changes in the usual experimental variables; but we find certain regularities within the test correlation matrices suggesting recurrent and meaningful factors which repeatedly appear throughout the different batches of data. In the treatment in terms of quasi-properties, on the other hand, the pointer readings are themselves invariant to changes in all the defining variables after multiplication by functions of those variables whose nature is defined by the quasi-property. It therefore seems unlikely that the two methods will generally disclose the same type of underlying entities.

---

[1] G. W. Scott Blair and J. E. Caffyn, *Journal of Scientific Instruments*, 1942, **19**, 88 ; G. W. Scott Blair and B. C. Veinoglou, loc. cit.

[2] G. Halsey, *Journal of Applied Physics*, 1947, **18**, 1072

In conclusion, it may be well to forestall a possible criticism of the use of quasi-properties and perhaps even of the factoring of data from empirical tests. It is sometimes suggested that the specification of materials in terms of any kind of magnitudes which are not completely invariant at least to changes in the variables involved in their definition, while possibly justified for purely *ad hoc* purposes, is of no interest to the theoretician.

If is doubtful if such an argument could find much support today either from physicists or psychometrists, but it has surely been dealt its death blow by the work of Dingle as exemplified both in the paper already quoted and in the earlier members of the series of papers of which it forms a part.

If, even in the simplest cases, ‘ physical properties ’ have only a limited invariance, there can be no objection to our pursuing such partial invariance as may be found in the case of highly complex materials.

6 *Summary*

The classical concept of the ‘ physical property ’ is examined and three alternative methods of classifying data for systems which do not show invariant physical properties are discussed.

The first method is analytical and accounts for behaviour in terms of additive combinations of finite or infinite numbers of simple prototype units, though in recent times complex units have also sometimes been postulated. This treatment is exemplified by exponential functions.

The second method establishes a certain measure of invariance in groupings of correlation coefficients between entirely empirical tests and finds its expression in matrix algebra.

The third method seeks invariant quantities at the expense of altering the functional relationships between the variables defining the 'property' in accordance with the nature of the material and the testing process and usually leads to power-law and fractional differential equations.

The last two methods are compared and shown to have much in common, but it is concluded that only in trivial cases will they lead to the same expressions of behaviour.[1]

G. W. Scott Blair

---

[1] The author is indebted to Dr. R. Harper for helpful discussions.

Natural Philosophy of Cause and Chance, Max Born, Clarendon Press, Oxford, 1949. Pp. viii + 215. 17s. 6d.

THIS book contains the substance of the Waynflete Lectures which Professor Born delivered at Oxford in 1948. The text has been somewhat modified and an Appendix, almost as long as the book itself, has been added in order to give mathematical proofs of the theorems referred to.

Although this *Journal* is not concerned with theoretical physics as such, it is impossible to comment on the book without a word of admiration of the way in which a vast part of the field of modern physics has been surveyed. Mechanics, thermodynamics, kinetic theory, relativity, quantum theory—the fundamental principles of these sciences are presented in a manner which the reader who is not ignorant of mathematics will find invaluable for the insight which it gives not only into the derivation of the various results but also into their relations with one another. One cannot but be impressed by Professor Born's mastery of his subject and his infallible eye for selecting the essential from the incidental. I know of no conspectus of theoretical physics more enlightening than this.

Our main concern, however, is with the philosophy of the book, and this, though no less thought-provoking, gives more opportunity for controversy, for it seems that the precision of thought so characteristic of the physics has not been bestowed in equal measure on the implications drawn therefrom. The whole discussion gathers round three ideas—causality, determinism, chance. Yet although causality and determinism are specifically defined, chance is not ; we have to discover, by inference from isolated statements, what it means. Thus on p. 1 we are told that it means the opposite of ' the idea of necessity in the relation of events '—' complete randomness '—but this is immediately qualified by the remark that there are ' laws of chance,' so that it cannot be the opposite of the idea of necessity. On p. 47 the fact that science makes use of the theory of experimental errors is held to show ' that the conception of chance enters into the very first steps of scientific activity, in virtue of the fact that no observation is absolutely correct.' ' I think,' adds Professor Born, ' chance is a more fundamental conception than causality ; for whether in a concrete case a cause-effect relation holds or not can only be judged by applying the laws of chance to the observations.' This seems to imply that chance is related to ignorance of the extent of our imperfections, which is a notion quite different from that of absence of necessity in the relation of events. How the two ideas are connected we are not told.

Again the distinction between causality and determinism leaves the careful reader perplexed. 'I think,' writes Professor Born, 'one should not identify causality and determinism. The latter refers to rules which allow one to predict from the knowledge of an event *A* the occurrence of an event *B* (and vice versa) but without the idea that there is a physical timeless (and spaceless) link between all things of the kind *A* and all things of the kind *B*. I prefer to use the expression "causality" mainly for this timeless dependence.' As an example of determinism without causality the connection of a time-table with the arrival of a train is cited : you can predict the train's arrival from the time-table, but you do not regard the time-table as its cause. This seems clear enough, and suggests that causality includes determinism but supplements it by the idea of an objective link or law connecting the events considered. But formal definitions are immediately given in which this distinction seems to be wiped out entirely : they are as follows : '*Determinism* postulates that events at different times are connected by laws in such a way that predictions of unknown situations (past or future) can be made.' '*Causality* postulates that there are laws by which the occurrence of an entity *B* of a certain class depends on the occurrence of an entity *A* of another class, where the word "entity" means any physical object, phenomenon, situation, or event. *A* is called the cause, *B* the effect.' It is difficult to see any difference of meaning in these two statements : both imply that pairs of events are connected by laws of such a kind that, if you know the laws, you can infer one event from the other. On p. 17 the situation becomes still worse. So far from causality including determinism, or being identical with it, we are told there that in quantum theory 'determinism is abandoned' but 'causality is not affected.'

Now I do not for one moment think that Professor Born's ideas on the subject are so indefinite as these statements would suggest, but the reader has to infer those ideas from the statements, and I think that unless he already has some notion of the matter he will fail to do so. Why so clear a thinker can express himself so confusedly is a matter of opinion, and my opinion is that it is because Professor Born has not thoroughly liberated his philosophical, as distinct from his scientific, thinking from conceptions appropriate only to a pre-relativity, pre-quantum outlook. 'Quantum mechanics,' he says (p. 100), 'does not describe an objective state in an independent external world, but the aspect of this world gained by considering it from a certain subjective standpoint, or with certain experimental means and arrangements.' The first part of this statement is wholly consistent with physical practice ; the second is a longing, lingering look behind at the 'independent external world,' which in fact must be left out of the description altogether if one is to understand what is being done. It is impossible to give an account of modern physics free from confusion without abandoning *completely* the attitude of one describing an objective world and assuming that of one investigating the possibilities of experience. Instead of 'the aspect of this world . . .' Professor Born should have written 'the relations between concepts formed to represent the connectedness of our experience.' The question 'Is an event determined ?' then means simply, 'Can you determine it ?' Eclipses were undetermined before they were understood ; the weather is (mainly) undetermined now. The legitimate impulse that prompts the false question whether there is an objective 'law' that governs it is represented by the true question, 'Will we ever be able to determine it ?' and that, though we cannot now give an answer, raises no hopeless metaphysical problems.

I think that if Professor Born would bring his philosophical notions into line with his physics he would simplify his view of things very greatly. An interesting example is his discussion of the 'reality' of the electron. As he rightly indicates, the concept of reality varies from one person to another, and he chooses to define it (p. 103) as that which 'presupposes that our sense impressions are not a permanent hallucination, but the indications of, or signals from, an external world which exists independently of us. Although these signals change and move in a most bewildering way, we are aware of objects with invariant properties. The set of these invariants of our sense impressions is the physical reality which our mind constructs in a perfectly unconscious way. This chair here looks different with each movement of my head, each twinkle of my eye, yet I perceive it as the same chair. Science is nothing else than the endeavour to construct these invariants where they are not obvious. . . . From this standpoint I maintain that the particles are real, as they represent invariants of observation. We believe in the "existence" of the electron because it has a definite charge e and a definite mass m and a definite spin s ; that means in whatever circumstances and experimental conditions you observe an effect which theory ascribes to the presence of electrons you find for these quantities, e, m, s, the same numerical values.'

It appears, then, that our subjective apprehension of the world tells us something trustworthy about the world when it enables us to 'construct' invariants of observation. One might, however, object that the position and momentum of an electron are not 'invariants of observation.' Professor Born admits this, and says, 'Though an electron does not behave like a grain of sand in every respect, it has enough invariant properties to be regarded as just as real.' But electrons obey Fermi-Dirac statistics, which means that they cannot be counted like ordinary 'real' objects. What sort of 'real' thing is it that has no individuality ? Might not one just as well say that it *lacks* enough invariant properties to be regarded as *unreal* ? But in fact the whole question is an idle one. Leave the 'real independent objective world' out of the picture, and the question does not arise, and nobody is a penny the worse.

The same change of attitude would also reduce, though not altogether destroy, the divergence between Born and Einstein which is touched on in the last chapter. 'You believe in the dice-playing god, and I in the perfect rule of law in a world of something objectively existing,' wrote Einstein; and Born replies, 'We mortals have to play dice anyhow if we wish to deal with atomic systems.' But in fact what objectively exist are our experiences, and there is no reason to doubt that they exhibit a perfect rule of law. Atomic systems are part of the logical machinery for describing the relations between experiences, and the dice you throw in order to deal with them are another part; you don't reach experience until you have thrown the dice and worked out the statistical results. If I find a law which tells me that the result of a certain measurement will be 6, and it is 6, I have no right to proclaim a fundamental indeterminacy in nature because I am unable to tell whether the 6 is 'really' $5 + 1$ or $4 + 2$ or $3 + 3$.

I have no space to refer to many aspects of this exceedingly interesting book which it is tempting to discuss, so I will conclude with a brief remark on only one of them. Professor Born considers at length the well-known difference between the reversible laws of mechanics and the irreversible laws of thermodynamics, and reaches (p. 72) the interesting conclusion that 'irreversibility is a consequence of the explicit introduction of ignorance into the fundamental laws.' There is a sense, and an important sense, in which this is true, but it is not the whole truth. Clausius's famous dictum was based not on statistical considerations but ultimately on the fact of experience that heat naturally tends to pass from hot bodies to cold and not in the opposite direction. And when we look at the *bare facts* of mechanics we see again the same one-way tendency, in spite of the reversible laws: bodies (in our region of space at any rate) tend to move towards and not away from one another. In both departments of our experience the *phenomena* show a one-way tendency; it would be natural, therefore, to expect that both sets of *laws* would be irreversible. This would seem to call for an explanation of the reversibility of the mechanical laws rather than for that of the irreversibility of the thermodynamical ones.

HERBERT DINGLE

The Power and Limits of Science: A Philosophical Study, E. F. Caldin, Chapman & Hall Ltd., London, 1949. Pp. x + 196. 12s. 6d.

THIS is a book written by a chemist who is interested in philosophy and who believes that there are important metaphysical presuppositions of natural science.

In Part I Mr Caldin outlines the method and scope of physics. He introduces the reader to its subject matter, the kind of knowledge to which it leads, and to inductive and hypothetical methods; he is concerned to bring out limitations of scientific explanation and that a philosophy cannot be based on physics alone. Part II is philosophical. Laws of nature cannot be deduced from observations. We have to ' interpret ' our data as signs in order to form a general law, which depends upon analogy ; and to justify this the author assumes Keynes's principle of Limitations of Independent Variety as an expression of the kind of order that must be supposed to exist in nature. He considers that scientific progress, which is achieved by better and better approximations, presupposes such a principle, but that grounds for believing it must lie in metaphysics. In this field he relies on ' realist ' theory of knowledge, widened by a treatment of beliefs that support and correct one another, and he holds that metaphysical conclusions can be obtained by a rational interpretation of experience. This is his main thesis. In Part III he considers other matters. He compares the beauty of science with other forms of beauty and suggests that different forms of it are characterised by types of ' variety-in-unity ' ; he concludes by discussing the relation of science to the good life and to society. He adds an appendix on the cosmological argument for a first cause, which is well stated—without reference, however, to the usual criticisms.

Mr Caldin has good things to say about belief and interpretation. But his central contention about metaphysical presuppositions attainable by rational interpretations is presented without analysis. Thus he develops in a traditional way the conceptions of the potential and actual to account for change, adds to it Aristotle's doctrine of the four causes, and claims that an Aristotelian essentialist definition of man, for instance, can be reached by reflecting on our own operations ; and, with Aquinas, he argues from potentiality and actuality to a first cause. No new support for these ancient doctrines is offered.

The book seems to have been written to introduce the general reader to the possibility of metaphysical implications of science, and to awaken him to relationships between science and aesthetics, morals, and society. Such a reader should be warned that the beliefs expressed are far from typifying the current climate of scientific or philosophical thought. The metaphysical outlook conveyed is roughly that of Aristotle or Aquinas or a blend of the two. The points of view of the influential present-day schools of empirical philosophy are not discussed.

J. O. Wisdom

The Foundation of Human Thought, Fr Vinding Kruse, Oxford University Press, London (Einar Munksgaard, Copenhagen), 1949. Pp. ix + 404. 30s.

Many thoughtful minds are deeply exercised by, and indeed apprehensive of, the world-wide confusion existing at the present time between science—

broadly so-called—and axiology. No doubt some degree of tension between these two subjects has existed for several centuries, whether openly recognised or not. Nevertheless, our own generation has become acutely aware of it, for reasons with which most people are familiar. It is not easy, however, to put one's finger accurately upon the crucial point ; the area in question is somewhat too diffuse, and includes, it seems, the advance of applied science and the deterioration of ethical standards. These things are scarcely matters of dispute : in fact they are all too evident ; certainly they demand analysis if the task can be undertaken objectively and dispassionately. The scientific method is said to be imperative : indeed it is, but is not in itself altogether sufficient. What is needed in addition is nothing less than charity, combined with a determination to be constructive rather than to act the fault-finder, to encourage where possible, rather than to blame. It is too much to hope that there will be universal agreement that these problems of our age can be solved or even ' softened up ' along these general lines. But there should be gratitude to any author courageous enough to face such issues squarely, and to share his thoughts with others in a similar state of perplexity.

Thus, a warm welcome can be extended to this book, the work of a professor of jurisprudence in the University of Copenhagen. The subtitle is significant, ' The problem of Science and Ethics,' for that is what the author is manifestly concerned to discuss. Originally, the script appeared in Danish (in 1942), and has since then been translated into English to form the present volume. Denmark eight years ago must indeed have provided a setting likely to exhibit profound philosophic difficulties to a mind of Dr Kruse's calibre : maybe good has come out of evil in thus forcing them to the surface. For the rest, the English version now before us contains an altogether abnormal number of misprints (they tend to occur in batches, some of them quite fantastic), which a second impression will doubtless remove. This said, they will not be referred to again ; few, if any of them, mislead, but they are like fine dust-particles sticking to a majestic picture.

In this essay, we shall not attempt to traverse those pages with a microscope, focusing comment or criticism here and there with accompanying limitation of field. Rather is it the purpose to follow the broad sweep of the writer's meditations, and to concentrate their essence in the hope that in this way, the pattern of his thought may be discerned.

Professor Kruse's book falls into three main parts—one dealing with the problem of ethical values, another with the nature of science and human knowledge, and the third with the question of ethical and religious evaluations.

So far as the first two are concerned, the broad conclusion is reached that ethics, as commonly understood, cannot be founded on knowledge, because the nature of knowledge itself has never been fundamentally clarified. Thus, the author recommends recasting the whole subject, and taking up a more empirical attitnde. His reason for doing this is the conviction that ethics can indeed be a science if its claims are made sufficiently modest, and its field correspondingly narrowed. Further, it is maintained that ethics has hitherto failed for two basic reasons : (*a*) because it has not sifted the fundamental problem of knowledge (which is tantamount to saying that it has condoned the general inability to do so) ; and (*b*) because it has omitted to take advantage of the present vast store of practical experience available, especially in the domains of medicine and jurisprudence.

In thus bringing ethics within the scope of the experimental sciences, Professor Kruse remarks that, in the last resort, its goal is the same as theirs, namely, the preservation of mankind. The bridge between this purposeful aim and the theoretical sciences discussed elsewhere is that *experiment* adds a fourth postulate to those commonly accepted : (i) that something *is*, (ii) that something *has been* and (iii) that something *will be*. Manifestly, deliberate interference by mankind is closer to the last of these than to the other two, though epistemologically it makes no odds whether in fact the intervention is natural or human. Nevertheless this emphasis upon the future is not to be taken as the third theorem of trivalent logic (possibility), but rather in the sense that something must necessarily, or very likely, happen. That ethics has attempted too much is the main reason why, on this view, it has achieved so little. If it is to recover, it must be by becoming an applied science.

How this may come about is obviously subject to severe limitations, and these Professor Kruse accepts. For example, we cannot in general be confronted with more than the hypothetical imperative—' *If* you want such-and-such a benefit, you must act in a certain way (e.g. by restraint and self-control).' The categorical imperative ' you must,' without the conditional, may be valuable as a support to the weaker form, which is the most that science can legitimately press. Kant tried, as is well known, to justify the categorical command scientifically, and in this was unsuccessful. It belongs to a plane beyond human faculties of cognition.

In the present context, however, a curious hint is thrown out. It is that theologians should cease to bother about certain abstruse questions, particularly linguistic, and concern themselves more directly with the abnormalities of the souls of men. It is very difficult to follow this suggestion, especially coming from so clear-headed an advocate as Professor Kruse. By definition, a theologian is a person who discourses about God, inclduing maybe what ' proofs ' or ' ways ' are believed to be available relating to His existence. And for this purpose, historical research, linguistic and otherwise, is one of his tools. So are many other branches of knowledge. But *qua* theologian, it is no business of his what men do to each other, and how they order their affairs. His own particular investigation may well persuade them to apply their hearts unto wisdom, but that is consequential rather than his primary, or even his principal, object.

Historically, the existence of many and strange gods in the mind of primitive man dictated, as it was bound to do, the rules of morality and justice. Against such a background, the revolution in outlook brought about by Thales of Miletus stands as a turning-point in the whole of human endeavour, even though it was to lead, for a period at least, to profound scepticism.

The intense individualism of Socrates illustrates the magnitude of the problem, and how easy it was, even in his day, to claim universal good, and simultaneously to enthrone authority. The assertion that virtue is knowing takes little root until we have been told what knowing is. This weakness is fastened upon by Professor Kruse as the source of all subsequent frustration, though it is legitimate to doubt whether countless thinkers since then have evaded the question so effectively as Professor Kruse implies. Centuries later came the view expressed quite definitely by Pomponazzi that ethics was independent of religion. The significance of this well-known doctrine has been profound ; its special bearing, however, upon the particular view expounded in this book becomes clear as the argument gathers momentum. We are not provided at any time with a final answer as to *why* a certain (in fact admirable) code of behaviour is to be striven after, except that it is consonant with incontrovertible principles of justice. That is, perhaps inevitably, as far as a scientific investigation of the whole subject can take us. It amounts to the simple precept that men shall not harm one another.

Professor Kruse has a great deal more to say on this score, which is both illuminating and profound, but which does not seem to extend the salient any further into the unknown, or at least into the unexperienced. The Queen of Sheba, we are told elsewhere, with all her *entourage* found much more than she had expected to find at the Court of King Solomon, and yet a Greater than Solomon was to come.

The middle chapters are very closely reasoned, and present a critical review of the English empiricists, and of Kant. Whilst it must be conceded that this inquiry is conducted with distinction, the upshot is that, after all the effort of ratiocination we find ourselves with something incapable of proof, the presence of elements of cognition which may not be sensations only. This statement appears to be of somewhat doubtful validity in itself, and is assuredly no more amenable to rigorous demonstration than the concept it is intended to replace.

Yet something very valuable does emerge, and that is the fundamental part played by likeness and difference as the condition of all knowledge. The mind is indeed capable of holding impressions after the external stimulus has ceased. That much is implicit. But it would be encouraging to say the least, if we could be told how this comes about, not from a merely mechanical angle, but maybe from an aspect in harmony with *Gestalt* psychology. The ' trace ' theory, once so attractive, seems unable to ' make the grade ' in the demand for an explanation of mnemic effects. Nevertheless it is one of the advantages of the present system that some ' apparatus ' suitable for the responsible duty of distinguishing—and of maintaining an ' afterglow '—is so clearly required.

Naturally, it is not being asserted that *because* Professor Kruse has decided to use the scientific method in his attempted synthesis of science and ethics that *therefore* he is under an obligation to offer purely scientific explanations of all the phenomena he invokes. Physico-chemical processes are obviously at work, though something more may have to be tolerated, something ' functional ' in the sense of E. S. Russell, even if it brings in a vestige of teleology. Without at least an element of this character, it is hard to see how ethics is going to be rescued from being merely a mechanical product, beneficial or not, as the machine may determine. If the reviewer understands Professor Kruse aright, this is not what he desires : his impressive reserve and the statuesque quality of his mind combine to prevent him from ' letting himself go,' as indeed is understandable on judicial principles. All of which is but a way of asking the basic question : ' Is the scientific method of necessity the correct one to use for the purpose after all ?' Perhaps it is, so long as it is allowable to discern his path still winding ahead onwards and upwards when reason has exhausted herself in the noblest of her crusades. From what the author says of the great religions of the world, one would infer that the *Deus philosophorum* is not the end of his journey.

Probably the most valuable suggestion which this book has to make concerning our attitude to fundamental questions is that called ' feeling-ahead.' It is also called ' willing-ahead,' thus implying that feeling includes an act of volition. Now this groping or tentative property has made its appearance not only at almost every epoch of human progress, but in every type of progress. If it be considered merely empirical, then experience seems reluctant to abandon it even when its trial-and-error character is fully appreciated. Axiologically it proves to be significant, and for the reason that it connotes a basic humility. Epistemologically, historically, aesthetically, creative minds have used it, in contrast to the ' once-for-all ' and ' streamlined ' methods of the imitator and the second-rate. Artistically, the great painters of classical pictures must have followed an analogous technique : radiographic analysis reveals many instances of line and contour which indicate unmistakably the process of feeling the way ' here a little and there a little.' In common with all purposeful processes, *Gestalt* psychology has included systems of this kind within a general integrative scheme, amenable in principle to the theory of stationary states of minimal energy.

The formal array which the author adopts for classifying what he calls ' the descriptive or theoretical sciences ' on the one hand, and ' the experimental, evaluating or practical (applied) sciences ' on the other is suggestive and neat. There is, in essentials, a one-to-one relationship between individual entries in either group. For example, psychology, history (including archaeology), philology, psychiatry, in the ' theoretical ' category are reflective into psycho-therapy (mental hygiene), individual ethics, aesthetics, in the ' applied ' category. And so on. In general, the latter activities are the result of a greater or lesser degree of interference on the part of mankind with the law-governed relations which characterise them : by contrast, likeness and difference suffice for the former.

From this brief survey, it will be obvious that Professor Kruse has contributed something of quite unusual value to epistemology and ethics. This is not merely or even mainly, because he writes with a wide erudition and discernment ; it is equally on account of the transparent sincerity of his tenets, and the engaging quality of his gentleness. The argument is intensely felt throughout, yet with no suggestion of ' breathing out threatenings and slaughter.' He takes all knowledge as his universe of discourse, but in so doing never permits himself to forget that his theme is not *super flumina*, but *de profundis*.

IAN RAWLINS

BOOKS RECEIVED FOR REVIEW

Inclusion of books in this list does not preclude their being reviewed in later issues

R. C. Brimley, *Organum*, Deighton, Bell & Co. Ltd., Cambridge, 1950, pp. 63, 3s.

W. I. B. Beveridge, *The Art of Scientific Investigation*, Wm. Heinemann, Ltd., London, 1950, pp. xii + 171, 10s. 6d.

G. W. Scott Blair, *Measurement of Mind and Matter*, Dennis Dobson Ltd., London, 1950, pp. 115, 9s. 6d.

A. Leslie Banks, *Man and His Environment*, Cambridge University Press, 1950, pp. 15, 1s. 6d.

G. Burniston Brown, *Science, its Method and its Philosophy*, George Allen & Unwin Ltd., London, 1950, pp. 189, 15s.

K. G. Collier, *The Science of Humanity*, Thomas Nelson & Sons Ltd., Edinburgh, 1950, pp. xx + 339, 12s. 6d.

Vergilius Ferm, *A History of Philosophical Systems*, Philosophical Library, Inc., New York, 1950, pp. xiv + 642, $6.00.

REVIEWS

I. J. Good, *Probability and the Weighing of Evidence*, Charles Griffin & Co. Ltd., London, 1950, pp. viii + 119, 16s.

Chauncey D. Leake and P. Romanell, *Can We Agree ?* University of Texas Press, 1950, pp. xiii + 110.

Arnold Lunn, *The Revolt Against Reason*, Eyre & Spottiswoode (Publishers) Ltd., London, 1950, pp. 252, 15s.

I. Manton, *Problems of Cytology and Evolution in the Pteridophyta*, Cambridge University Press, 1950, pp. x + 316, 45s.

Edward A. Maziarz, *The Philosophy of Mathematics*, Philosophical Library, New York, 1950, pp. viii + 286, $4.00.

W. D. Ross, *Aristotle's Prior and Posterior Analytics*, Clarendon Press, Oxford, 1949, pp. x + 690, 42s.

G. Sommerhoff, *Analytical Biology*, Oxford University Press, 1950, pp. viii + 207, 17s. 6d.

Charles Singer, *A History of Biology*, Revised Edition, H. K. Lewis & Co. Ltd., London, 1950, pp. xxv + 579, 35s.

Stephen Toulmin, *The Place of Reason in Ethics*, Cambridge University Press, 1950, pp. xiv + 228, 16s.

Barbara Wootton, *Testament for Social Science*, George Allen & Unwin Ltd., London, 1950, pp. vii + 197, 15s.

OBITUARY

PROFESSOR E. A. MILNE, F.R.S.

It is with the deepest regret that we have to announce that the Philosophy of Science Group of the British Society for the History of Science has lost a distinguished member of its Editorial Board by the sudden death on 21st September 1950, while on a visit to Dublin, of E. A. Milne, F.R.S., Rouse Ball Professor of Mathematics in the University of Oxford and Fellow of Wadham College.

During his comparatively short life of fifty-four years, Edward Arthur Milne, who was born on 14th February 1896, wrote so many memoirs of the highest quality and originality on a wide variety of subjects that he must be ranked among the leading men of science of his generation. After attaining international fame as a theoretical investigator, first of the properties of stellar atmospheres and then of stellar structure, Milne's name in recent years has been chiefly associated with particular theories of relativity and cosmology. In this way, beginning as a mathematical physicist whose field of study was highly specialised, his outlook gradually became more synoptic and far-reaching. Like Michael Faraday a century before him, he loved the term *natural philosopher*, and preferred that designation to all others.

In his inaugural lecture at Oxford, delivered in 1929, on *The Aims of Mathematical Physics*, Milne gave clear expression to his basic conceptions of natural philosophy and in particular to the role which he believed that mathematical theorems should play in the study of natural phenomena, ideas which he subsequently developed on a far grander scale in his writings on Kinematic Relativity. One of his greatest ambitions was to do for dynamical, gravitational, electromagnetic and other phenomena what the Greek geometers did for purely spatial phenomena. The measure of his success in achieving this object is overshadowed by the novelty and grandeur of the whole conception. As with so many other great thinkers, particularly philosophers, the abiding significance of Milne's work is likely to depend on the questions which he was the first to raise, irrespective of the merit of the particular answers which he himself provided.

A great philosopher of nature is gone from us and we mourn his untimely passing.